# Who Counts as a Semite? Etymology, Linguistics, and the Y-Chromosome

**Noël Alejandro Bota, J.D.** · Research and drafting assisted by Claude (Anthropic)

## The question splits in three

"Who is a Semite?" is not one question. It is three, and they have different answers.

1. **Who speaks a Semitic language?** Answerable, uncontroversial, and about languages rather than people.
2. **Who descends from the people who first spoke Proto-Semitic?** Answerable in principle, but the answer turns out to be nearly everyone in the relevant region, for reasons of arithmetic rather than history.
3. **Which paternal lineage "founded" Semitic?** The question most often asked in genetic-genealogy circles, and the one this document argues is malformed.

Most disagreement about the term comes from sliding between these without noticing. The document below works through the etymology, the linguistic dating, the ancient DNA, and one specific thesis — that Y-haplogroup E-M35 represents the aboriginal Afro-Asiatic lineage and therefore the "real" Semites by descent.

## The word is a 1781 filing label

"Semite" is a European neologism roughly 245 years old. No ancient person ever used it of themselves.

The chain: Hebrew **שֵׁם** (*šēm*) → Greek **Σήμ** (*Sēm*) → Late Latin **Sem** → Modern Latin **Semita** → German *semitisch* / French *Sémite* → English *Semite*.

Two points of pedantry matter.

שֵׁם is not an opaque proper noun. It is the ordinary Hebrew common noun for "name," "renown," "reputation." Shem is literally Name. In Genesis 9–10 he functions as an eponym — a personified "ethnolinguistic bracket — not a person with a biography. Eponyms do not have haplogroups.

**The coinage is documented and recent.** The first known use is in August Ludwig von Schlözer's article on the Chaldeans in Eichhorn's *Repertorium*, vol. 8, p. 161 (1781); Eichhorn later made a competing claim to the invention in his *Allgemeine Bibliothek* (1794). The English noun "Semite" appears only in 1847, meaning a Jew, Arab, Assyrian, or Aramaean, as a back-formation from "Semitic" or from French *Sémite* (1845).

The category was invented to file languages. It was borrowed from a genealogy that was never about languages.



*© 2026 Noël Alejandro Bota, J.D. — Educational, historical, and genealogical research use only (Fair Use). Drafting and editing assistance with Claude-AI. Verify all claims against primary sources.*

## Genesis 10 classifies Hebrew as Hamitic

The source text contradicts the category built on it, and it does so immediately.

Genesis 10 assigns to Shem: Elam, Asshur, Arpachshad, Lud, Aram.

| Biblical assignment | Actual linguistic affiliation | Problem |
|---|---|---|
| Elam → Shem | Elamite is a **language isolate** | Not Semitic, not even Afro-Asiatic |
| Lud → Shem | Lydian is **Anatolian Indo-European** (if Lud = Lydia) | Wrong family entirely |
| Canaan → **Ham** | Canaanite, Phoenician, Moabite, **Hebrew** | The closest relatives of Hebrew are Hamitic in the text |
| Asshur, Aram → Shem | Akkadian, Aramaic | These two work |

Note what the text does **not** say. It does not place the Hebrews under Ham. Shem's line runs Arpachshad → Shelah → Eber → Peleg → ... → Terah → Abram, and Genesis 10:21 pointedly calls Shem *the father of all the children of Eber* — עֵבֶר, the eponym from which עִבְרִי (*ʿivrî*, "Hebrew") is traditionally derived. The Hebrews are Shemite **in descent**, unambiguously.

What is Hamitic in the text is the **language**, not the people. That distinction is the single most interesting thing in Genesis 10, and it is explored in the next section.

The Table of Nations sorts by geopolitical alignment and imperial relations in the first millennium BCE, not by linguistic descent. Schlözer grafted a linguistic category onto a political one, and everything downstream inherits the error. Anyone asking "who descends from Shem" is asking a question whose source text would exclude Hebrew speakers and include Elamites.

## The Bible records a language shift — and it is the one that names the category

The text separates descent from speech, and then tells us the Hebrews changed languages. This is the most thought-provoking thing in the whole source, and it anticipates the central finding of this document by about 2,700 years.

**Shemite people.** Abraham descends from Shem through Arpachshad and Eber. Not in dispute within the text.

**Canaanite language.** Isaiah 19:18 calls the tongue of Judah **שְׂפַת כְּנַעַן** (*śəpaṯ Kənaʿan*) — literally the *lip of Canaan*. The prophet's own term for the language of Shem's chosen line is the language of Ham's son. Modern linguistics agrees: Hebrew is a Canaanite language in the Northwest Semitic group, closer to Phoenician and Moabite than to Aramaic. Koert van Bekkum takes the phrase as the title of his survey of the problem, "The 'Language of Canaan': Ancient Israel's History and the Origins of Hebrew" (2018), which traces how assumptions about Israel's history have shaped linguistic accounts of Hebrew's origins — and notes that the earliest Hebrew inscription, the Gezer calendar, dates only to the 10th century BCE.



*© 2026 Noël Alejandro Bota, J.D. — Educational, historical, and genealogical research use only (Fair Use). Drafting and editing assistance with Claude-AI. Verify all claims against primary sources.*

**And the text stages the shift explicitly.** In Genesis 31:47, Laban names the stone heap in Aramaic (*Yəḡar Śāhăḏûṯā'*) while Jacob names the same heap in Hebrew (*Gal'ēḏ*) — a bilingual moment marking the point where the family's Mesopotamian branch and its Canaanite branch diverge in speech. Deuteronomy 26:5 has the Israelite confess that his father was a wandering (or perishing) **Aramean**. The patriarchal line is depicted as coming out of an Aramaic-speaking milieu in Ur and Harran, entering Canaan, and taking up the local tongue.

**And a Semitist stated it plainly.** Jan P. Lettinga's Biblical Hebrew grammar, reworked for Brill in 1962, classified Hebrew as Northwest Semitic — abandoning the older characterization of it as a *Mischsprache* — and held that Hebrew developed from the Canaanite dialect the Israelite tribes encountered on entering the land. They adopted that dialect, he wrote, and "almost entirely gave up their native language," which he took to be closely related to Old Aramaic.

The revision history matters as much as the claim. In the 1976 edition Lettinga added a reference to Deuteronomy 26:5, but **softened** his wording on Old Aramaic from "closely related" to "probably related" — most likely because the oldest Old Aramaic inscriptions date only to the **9th century BCE**. That is the same anachronism caveat this document raises about Genesis 31:47, and a leading Semitist hedged his own formulation for exactly that reason. Koert van Bekkum discusses the passage and its revisions in "The 'Language of Canaan': Ancient Israel's History and the Origins of Hebrew" (2018), noting that the origins of Hebrew and its place among the Semitic languages have been contested since the rise of modern historical linguistics.

## The implication

If the Bible's own account is taken at face value, **Hebrew is an acquired language for the Shemite line, not an ancestral one.** What the Hebrews spoke before Canaan was something else — Aramaic, or whatever the text imagines was spoken in Ur of the Chaldees.

And here the category eats itself. Schlözer's 1781 coinage was made for the Hebrew group; the word *Semitic* was built outward from Hebrew to Aramaic, Arabic, Akkadian and the rest. So the paradigm case on which the entire family name rests is, by the Bible's own testimony and by modern classification alike, **a Canaanite language that the children of Shem adopted from the children of Ham.**

The category "Semitic" is therefore named after a people who, in the source text, spoke someone else's language.

## Three caveats

1. This is internal to the narrative's logic. The patriarchal accounts are not history, and no independent attestation of Abraham exists.
2. The Aramaic of Genesis 31:47 is anachronistic — Aramaic is attested as a distinct language only from roughly the 10th–9th centuries BCE. The verse tells us how the text's authors understood their own origins, not what anyone spoke in the second millennium.
3. Isaiah's phrase may be geographic rather than genealogical — "the language spoken in the land of Canaan" rather than "the language belonging to the lineage of Canaan." Commentators divide on this.



*© 2026 Noël Alejandro Bota, J.D. — Educational, historical, and genealogical research use only (Fair Use). Drafting and editing assistance with Claude-AI. Verify all claims against primary sources.*

### Why it matters to the genetic argument

Strip away the theology and the structure is exactly the pattern found at Peqi'in and among Chadic speakers: **a population whose descent runs one way and whose language runs another.** The Bible encodes a people who kept their genealogy from Shem and took their speech from Canaan.

Whoever compiled Genesis 10 and Isaiah 19 evidently did not think that speaking a language made you a descendant of its first speakers, or that descent obliged you to keep your ancestors' tongue. The modern argument that a Y-haplogroup can establish who is "really" a Semite asks the category to do something its own source text never asked of it.

## Why is Canaan under Ham? The Egyptian-province reading

There is a specific and well-supported scholarly answer to why a Semitic-speaking people got filed under Ham alongside Egypt, Cush and Punt: **Canaan was an Egyptian province.**

Throughout the Late Bronze Age (c. 1550–1200 BCE), Egypt ruled Canaan through a network of vassal city-states, with governors' residencies and garrisons at Beth Shean, Jaffa, Gaza, Lachish, Aphek and elsewhere. Lester Grabbe concludes that there were very few periods during the Late Bronze Age when Palestine was not firmly under Egyptian control.

Meindert Dijkstra puts the connection to Genesis 10 directly: knowledge of Canaan as a colony of Egypt lies disguised in the genealogical tradition that counts Canaan among the sons of Ham together with Egypt, Nubia and Punt. On this reading, **Genesis 10 is a political map of the Egyptian imperial sphere, remembered as a family tree.** Ham is not a race or a language group. Ham is Egypt and everything Egypt ruled.

A second, compatible reading attaches to the Curse of Canaan in Genesis 9: a land claim on a territory that had fallen under New Kingdom Egyptian rule has been proposed as the motive both for cursing Canaan specifically and for affiliating him with Ham by way of Egypt's dominion.

### The dissent

Israel Finkelstein rejects the premise that the Bible preserves any Bronze Age memory at all, stating that it has no recollection of the situation in the Bronze Age, no memory of the crisis at its end, and no memory of Egypt in Canaan. If he is right, the Egyptian-province explanation cannot be a surviving memory — though the grouping could still reflect a later author's geography rather than a preserved one. The dispute is live.

## Arphaxad and the Chaldeans: a tradition, not a datum

The line from Shem to Abraham runs through Arpachshad, and a long tradition connects him to the Chaldeans. The tradition is real. Its evidentiary value is weaker than it is usually presented.

**What is solid.** Josephus states it plainly in *Antiquities* 1.6.4: Arphaxad named the Arphaxadites, who are now called Chaldeans. The same passage has Shelah son of Arphaxad and Eber his son, *from whom they originally called the Jews Hebrews* — so the Eber etymology is explicit in Josephus too.



*© 2026 Noël Alejandro Bota, J.D. — Educational, historical, and genealogical research use only (Fair Use). Drafting and editing assistance with Claude-AI. Verify all claims against primary sources.*

**What is contested.** The etymology is not settled. Competing proposals include Umberto Cassuto's *Arapcha-Kesed*; an identification with Arrapḫa / Arrapachitis in Assyria; Schrader's reading as the coast or boundary of the Chaldeans; and Hommel's suggestion that it is an Egyptianized spelling of Ur of the Chaldees. These cannot all be right, and no consensus exists. Any presentation of *Arpha + Kshad* as the established derivation overstates the case.

**What complicates it further.** The Hebrew Bible supplies a *second* Chaldean eponym. Genesis 22:22 lists **Kesed** (כֶּשֶׂד) among the sons of Nahor, Abraham's brother, in Aram-Naharaim; Jubilees 11:7 has Ur son of Kesed build the city of Ur-Kasdim. So the text offers two unrelated routes to the Kaśdim, and they are not reconcilable.

**And the anachronism stands.** The Kaldu arrived in southern Mesopotamia in the early first millennium BCE, after earlier Sutean and Aramean waves, and are attested from roughly the 9th century. A second-millennium patriarch departing *Ur Kaśdim* is being described in first-millennium terms regardless of which etymology is preferred. Josephus, writing in the 1st century CE for a Greek-speaking audience, is reporting a tradition roughly two millennia downstream of the events it describes.

Note also that Josephus identifies Elam's descendants as the ancestors of the Persians — which is only half-correct, since Elamite is a language isolate and Persian is Indo-Iranian (albeit, modern research has proven substantial blending between these two cultures, as well as the Bactria-Margiana-Oxus-Civilization, or “BMAC”, for short). His Table is a 1st-century geography, not a linguistic genealogy.

### Vetting note on circulating summaries

Popular and AI-generated syntheses of this material get much of it right but reliably introduce three errors, each of which this document avoids:

1. **"The sons of Shem map out the great Semitic-speaking peoples."** False, and self-contradictory wherever it appears alongside the correct observation that Elam is a language isolate. Elam is not Semitic-speaking; Lud is unidentified. The list does not map the Semitic languages — that is the entire point of the mismatch.
2. **"Hebrew and Phoenician speakers could understand each other with relative ease."** Plausible but unverifiable. The Canaanite languages were closely related, but claims of specific mutual intelligibility between ancient speech communities have no direct evidence behind them.
3. **Fringe sourcing.** Searches on this topic surface Rudolph R. Windsor's *From Babylon to Timbuktu* and similar works arguing that the original Israelites were “Black Shemites”. These Afrocentric works are not academic sources, do not reflect scholarly consensus, and are not used here. The question of ancient Levantine phenotype is addressed by the ancient DNA literature cited in this document, not by that tradition.

## Non-Semitic layers in Aramaic — and the Elamite that isn't there

Aramaic's foreign vocabulary stratifies cleanly by period, and the stratification is itself an argument.

| Layer | Period of contact | Depth of penetration |
|---|---|---|



*© 2026 Noël Alejandro Bota, J.D. — Educational, historical, and genealogical research use only (Fair Use). Drafting and editing assistance with Claude-AI. Verify all claims against primary sources.*

| Layer | Period of contact | Depth of penetration |
|---|---|---|
| Greek | Hellenistic, post-332 BCE | A handful of terms (Daniel 3's instruments) |
| **Old Persian** | Achaemenid, post-539 BCE | **Deepest** — vocabulary, verbs, idioms |
| Akkadian (and Sumerian through it) | Early 1st millennium onward | Vocabulary plus grammar |
| Hurrian | Late Bronze Age | Real, but concentrated in Ugaritic |
| **Elamite** | — | **No attested layer** |

**The Iranian layer goes past vocabulary.** Beyond the familiar *dāt* (law), *gizbār* (treasurer), *ʾăḥašdarpan* (satrap), *pitgām* (decree) and *rāz* (secret), Old Persian loans were absorbed into the Aramaic verbal system and conjugated as native roots: *gnz* "to store," *bgn/pgn* "to cry for help," *bšqr* "to search out," *grb* "to seize," *hrzq* "to lock up," *rwz* "to be glad." Hybrid idioms pair the Aramaic verb *ʿbd* "to do" with an Old Persian noun — *gasta patigāma*, *handarza*. That is structural integration, not surface borrowing.

**Sumerian arrives only through Akkadian**, centuries after it died as a spoken language: *hêkal* "palace" < Akk. *ekallu* < Sum. *é.gal*; *sārîs* "court official" < *ša-rēši*; plus *ʾiggartāʾ* (letter), *kursyā* (chair), *peḥâ* (governor), *ʾarḏîḵlāʾ* (architect). Watson cautions that the *hêkal* chain is "uncertain but feasible" rather than settled, and that *kussû*/*keš-i* may have travelled from Semitic **into** Hurrian rather than the reverse.

**Akkadian influence reached the grammar.** Aramaic's relatively free word order is plausibly an Akkadian outcome, and the Šap̄ʿel stem — rare in Aramaic, standard in Akkadian — may itself be borrowed.

### The discrepancy

**There is no detectable Elamite stratum in Aramaic**, despite extensive administrative contact: the Persepolis Fortification tablets are Elamite documents carrying Aramaic epigraphs, so scribes worked in both daily. The Achaemenid-era borrowings nonetheless came through Old Persian.

This matters for reading Genesis 10. If the placement of Elam at the head of Shem's sons reflected any real linguistic relationship, some Elamite residue would be expected somewhere in the region's Semitic. None is attested. Elamite remains an isolate that borrowed and lent across the Zagros frontier without leaving a mark on Aramaic.

*Caveat:* this is a negative finding from a general survey, not from the specialist Elamite literature. Treat it as provisional.

## What Shem's peoples actually share is Hurrian

Ask what the sons of Shem have in common that is **not** Semitic, and a single answer emerges — and it is not a language family any of them belonged to. It is the non-Semitic, non-Indo-European population whose territory their territories overlap.



*© 2026 Noël Alejandro Bota, J.D. — Educational, historical, and genealogical research use only (Fair Use). Drafting and editing assistance with Claude-AI. Verify all claims against primary sources.*

| Shem's son | Identification | Hurrian connection |
|---|---|---|
| **Arpachshad** | Often linked to **Arrapḫa** (modern Kirkuk) | A **Hurrian city** — capital of the Hurrian kingdom of Arrapha, under Mitanni domination; Hurrians took it from Assyrian control at the start of the 2nd millennium and settled as the majority population |
| **Aram** | Aram-Naharaim, upper Mesopotamia | Overlaps the **Mitanni heartland**; the Khabur valley was the core of Hurrian lands for a millennium |
| **Asshur** | Assyria, upper Tigris | Sustained Hurrian contact; Arrapḫa was Assyrian, then Hurrian, then Assyrian again by the mid-2nd millennium |
| **Lud** | Lydia — or **Lubdu**, listed among the ancient cities of the Kirkuk region beside Arrapḫa and Nuzi | On the second identification, the same Hurrian zone |
| **Elam** | Susiana, southwestern Iran | The outlier: a language isolate with no shared substrate |

The Hurrian city list of the second millennium reads: **Arrapḫa, Harran, Kahat, Nuzi, Taidu, Washukanni.**

Two of those names matter here. **Arrapḫa** is the leading candidate for Arpachshad. **Harran** is where Terah's family settles and where Abraham's line remains (Genesis 11:31; 24:4; 27:43).

On the text's own genealogy and the leading etymology, **the Abrahamic line is routed through two Hurrian centres.**

### How deep the contact ran

This was not light borrowing. A **Hurrian–Akkadian creole**, known as Nuzi, was spoken at Arrapḫa. Wilfred Watson identifies roughly **80 possible Hurrian loanwords** in Ugaritic alone. One Hurrian term reaches Biblical Hebrew and Aramaic directly: **tartān**, the Assyrian commander-in-chief (2 Kings 18:17; Isaiah 20:1), from Hurrian *turtānu*.

Layered on top of Hurrian in Mitanni sits an **Indo-Aryan superstrate** — a thin elite stratum in royal names, gods and chariot vocabulary. The Hittite–Mitanni treaty invokes *Mitra*, *Varuna*, *Indra* and *Nasatya*; Kikkuli's horse-training manual counts in Indo-Aryan numerals (*aika*, *tera*, *panza*, *satta*, *na*). The numeral *aika* specifically locates the superstrate in Indo-Aryan proper rather than Iranian, which has *aiva*.

### Chronology: the two layers are not contemporary

This distinction matters and is easily blurred. **The Hurrian layer is old enough to be contemporary with any patriarchal horizon. The Indo-Aryan layer is not.**



*© 2026 Noël Alejandro Bota, J.D. — Educational, historical, and genealogical research use only (Fair Use). Drafting and editing assistance with Claude-AI. Verify all claims against primary sources.*

| Element | Date | Relative to a c. 2000 BCE Abraham |
|---|---|---|
| First Hurrian kingdom at Urkesh (Tell Mozan) | 3rd millennium BCE | **Earlier** |
| Tish-atal of Urkesh inscription | Start of 2nd millennium BCE | **Contemporary** |
| Hurrians take Arrapḫa from Assyrian control | Beginning of 2nd millennium BCE | **Contemporary** |
| Harran attested in Ur III texts as a temple of Sin (*é.ḫul.ḫul*) | c. 2000 BCE | **Contemporary** |
| Indo-Aryan elite enters the region | c. 17th century BCE | **~300 years later** |
| Kikkuli's horse manual | c. 1400 BCE | **~600 years later** |
| Hittite–Mitanni treaty (Mitra, Varuna, Indra, Nasatya) | c. 1380 BCE | **~620 years later** |
| Arrapḫa as a Hurrian kingdom under Mitanni domination | 15th–14th centuries BCE | **~550 years later** |

So the Hurrian presence in the Shem zone predates or coincides with the traditional Abraham horizon, while the Indo-Aryan superstrate and the Mitanni-era kingdom of Arrapḫa postdate it by roughly three to six centuries. A Terah migrating from Ur to Harran around 2000 BCE would have entered a region with Hurrians already in it — but no Mitra, no Indra, and no Kikkuli.

### Which Abraham date?

There is no fixed date, and the range is wide enough to change the answer:

- **c. 2166 BCE** (birth) / **2091 BCE** (departure from Harran) — conservative chronology reckoned from the Masoretic figures and 1 Kings 6:1
- **c. 2000–1900 BCE** — Albright, correlating Abraham with Amorite movements and MB I; the most widely cited range
- **c. 1812 BCE** — *Seder Olam Rabbah*, placing Abraham's birth at 1948 AM
- **c. 1400 BCE** — Rendsburg and others who weight the genealogies over the stated year-figures, placing the patriarchs in the Late Bronze Age
- **No historical patriarchal age at all** — Thompson and Van Seters, who argue the narratives reflect first-millennium conditions and are Iron Age creations

On the first three datings the Indo-Aryan material is anachronistic. On a Late Bronze date, it becomes roughly contemporary. On the Thompson–Van Seters view the question does not arise, since there is no patriarchal period to date.



*© 2026 Noël Alejandro Bota, J.D. — Educational, historical, and genealogical research use only (Fair Use). Drafting and editing assistance with Claude-AI. Verify all claims against primary sources.*

The *Sin* connection is worth noting independently: Harran was a centre of the moon-god cult, closely related to the cult of Nanna/Sin at Ur. That link holds at the c. 2000 BCE horizon regardless of which Ur is meant, and it does not depend on any Hurrian or Indo-Aryan element.

So the region Genesis assigns to Shem was, in the second millennium, a Hurrian-speaking zone with later Indo-Aryan ruling vocabulary, writing in Akkadian, over a Sumerian lexical substrate. **Four non-Semitic layers.**

### Three caveats, and they matter

1. **This is contact, not descent.** Hurro-Urartian is a family of its own, unrelated to Semitic. Shared Hurrian vocabulary indicates neighbours, not ancestors.
2. **Geographic contiguity explains it without any genealogy.** Peoples occupying adjacent territory share substrates. This supports reading Shem as a **geographic and political bloc** — the conclusion reached elsewhere in this document — rather than as a lineage or a language group.
3. **The Nuzi-parallels literature is contested.** Wayne Pitard's caution is the standard one: Late Bronze Age parallels to marriage, inheritance and family-religious customs in Genesis cannot show that those stories preserve second-millennium traditions, because most such customs continued into the first millennium when Genesis was written down. The Nuzi texts have a limited, largely illustrative function.

### The net effect (humoring a J1-association)

The non-Semitic material shared across Shem's peoples is real, substantial, and points to the Hurrian zone of the upper Tigris, the Khabur and the Zagros piedmont — the same region identified elsewhere in this document as the source of the Chalcolithic-Iranian-related ancestry and the J1-P58 patrilines that entered the Levant in the Bronze Age.

That convergence is genuine and worth noting. It is also exactly what a first-millennium author's geography would produce, and it remains contact rather than descent. It sharpens the picture of **where** Shem is. It says nothing about which patriline spoke Proto-Semitic.

## The names say Amorite, not Hurrian

The territorial argument above would predict Hurrian elements in the patriarchal onomasticon. There are none. The names are **West Semitic, specifically Amorite** — and the Hurrian names in the Hebrew Bible cluster somewhere else entirely.

### The patriarchal line

**Abram** = *ʾab-rām*, "the father is exalted," on the West Semitic root *rām-* "to be high" — distinct from East Semitic Akkadian *rāmu* "to love." Well attested among Amorite names, with *Abi-ramu* appearing on Babylonian contracts around 2000 BCE. Alan Millard concludes that the patriarchal names exemplify a stage in the history of **West Semitic** naming.



*© 2026 Noël Alejandro Bota, J.D. — Educational, historical, and genealogical research use only (Fair Use). Drafting and editing assistance with Claude-AI. Verify all claims against primary sources.*

**Jacob** has its counterpart in *Yaqub-Har*, a West Semitic name borne by a ruler of Egypt's Second Intermediate Period. The Mari Amorite onomasticon supplies parallel structures such as *Iaḫwi-kî-Adad* and *Iaḫwi-kî-El*.

**Peleg, Serug, Nahor and Terah** all appear as **city names in the Harran region**, which is why they are generally read as eponymous ancestors of those communities. A real finding about the tradition's geography — and not a Hurrian one.

### Where the Hurrian names actually are

| Figure | Reference | Hurrian element |
|---|---|---|
| **Araunah** / Ornan, the Jebusite | 2 Samuel 24; 1 Chronicles 21 | *ewri* "lord, governor"; described as non-Semitic, probably Hurrian, possibly a title — it occurs once with the definite article. The variants *ʾrwnh* and *ʾwrnh* match the two attested Hurrian forms, *erwi-ne* (Mesopotamian, as at Nuzi) and *ewri-ne* (western Syrian) |
| **Shamgar ben Anath** | Judges 3:31 | Possibly Hurrian *Šimiqari*, attested at Nuzi, from **Šimige** the Hurrian sun god — worshipped from Alalakh to Arrapḫa. Pfeiffer, Speiser and Maisler argued this; the lexica still call the etymology uncertain |
| **Horites** (*ḥōrî*) | Genesis 14, 36; Deuteronomy 2 | The gentilic for a people in Canaan and Transjordan before Israel. Identification with Bronze Age Hurrians is raised and genuinely debated |
| Hebron figures | — | Mazar proposed derivations from Hurrian *Šawuška-šarru* |
| **Abdi-Ḫeba** of Jerusalem | Amarna letters (extra-biblical) | *Ḫeba* / Ḫebat, the Hurrian goddess |

A recent proposal reads *ʾwryh* (Uriah the Hittite) as a Hebraized *ewri*. It comes from a single study and should not be leaned on.

### The distribution is the point

| Group | Onomastic profile |
|---|---|
| Patriarchs — Abram, Isaac, Jacob | **West Semitic / Amorite** |
| Terah genealogy — Peleg, Serug, Nahor | **Harran-region toponyms** |
| Jebusites, Horites, Judges-era figures | **Hurrian elements present** |



*© 2026 Noël Alejandro Bota, J.D. — Educational, historical, and genealogical research use only (Fair Use). Drafting and editing assistance with Claude-AI. Verify all claims against primary sources.*

The Hurrian onomastic layer in the Hebrew Bible sits on the **Canaan** side of the Ham/Shem divide, not the Shem side — the reverse of what the territorial argument predicts.

The Hurrian connection to the patriarchs was always a claim about **customs** — the Nuzi adoption, surrogacy and inheritance parallels — never about names. And that is the literature Pitard's caveat deflates, since those customs continued into the first millennium when Genesis was written down.

**A Hurrian neighbourhood: yes. A Hurrian Abraham: no.**

## Reconciling three findings that look contradictory

A reader arriving at this point has been handed three claims that seem to pull against each other. They do not, and seeing why is the clearest statement of this document's thesis.

| Finding | What it is a claim about |
|---|---|
| Shem's territory carries a Hurrian substrate | **Geography** — where the bloc sits |
| Hebrew is "the language of Canaan" | **Language** of the Shemite line |
| Patriarchal names are Amorite | **Onomastics** of that line |

No two conflict, because no two are about the same category. The Hurrian finding was never evidence about identity — it is a statement about the neighbourhood, and the Amorite names are the proof that the neighbourhood did not determine the lineage.

### Three mismatches, one pattern

Amorite is **Northwest Semitic** — the same branch as Canaanite and Aramaic. So the patriarchal names place Abraham's line in the same language family as Canaan, which Genesis files under **Ham**, and not in the Hurrian family of the territory Genesis files under **Shem**.

That is a third instance of the same mismatch, not a counterexample to it:

1. **Canaan** — filed under Ham, speaks Northwest Semitic
2. **Elam** — filed under Shem, speaks a language isolate
3. **Abraham's line** — filed under Shem, sitting in Hurrian-substrate territory, bearing Northwest Semitic names

Every time the Table of Nations is tested against language, the label and the language come apart — three times, in three different directions. The consistency of the failure is the finding.

### A refinement this forces

The language shift described earlier needs sharpening. Laban names the heap in Aramaic and Jacob in Hebrew (Genesis 31:47); Deuteronomy 26:5 confesses a wandering Aramean. That is a genuine shift — but it is a shift **within Northwest Semitic**, since Aramaic and Canaanite are sister branches of one group.



*© 2026 Noël Alejandro Bota, J.D. — Educational, historical, and genealogical research use only (Fair Use). Drafting and editing assistance with Claude-AI. Verify all claims against primary sources.*

The patriarchal line therefore never crossed a language-family boundary. It moved between dialects inside a family it already belonged to. Linguistically, Abraham's family was on the "Hamitic" side of the Bible's own ledger **before it ever reached Canaan**.

### Two clarifications

**Shem is not "Hamitic."** The text keeps descent and speech apart: the Hebrews are Shemite *in descent* while their language is Canaanite. Collapsing the two reproduces exactly the error that Schlözer's 1781 coinage introduced.

**The shared material among Shem's peoples is not all Hurrian.** They also share a **Sumerian substrate transmitted through Akkadian** — the universal Mesopotamian layer. Hurrian is the distinctive shared element; Sumerian-via-Akkadian is the broader one. Elam shares neither in any detectable way.

## Does the Sumerian layer point to Sumer?

No. It points to the **cuneiform world**, which is nearly everyone — and a feature shared by everyone identifies no one.

The Hurrian layer was informative because it is **regionally restricted** to the upper Tigris, Khabur and Zagros piedmont. Sumerian-via-Akkadian marks nothing comparable, because Akkadian was the lingua franca of the entire Near East: Egypt and Canaan corresponded in Akkadian at Amarna. A Sumerian lexical stratum establishes that a language sat inside the cuneiform sphere, alongside Hittites, Hurrians, Elamites, Egyptians and Canaanites.

This is structurally identical to the E-M35 problem set out earlier: **a feature too widespread and too old to distinguish a subgroup.**

### Three further problems

**It is dead-language transmission.** Sumerian ceased to be spoken around 2000–1800 BCE and survived as a scribal and liturgical register. Aramaic is first attested in the early first millennium. So *hêkal* and *sārîs* reached Aramaic from a frozen classical language through scribal tradition, not from contact with Sumerian speakers.

**The direction of borrowing runs both ways, and more than once assumed.** One study catalogues **369 Akkadian loanwords in Sumerian** spanning the pre-Ur III period through Hammurabi. Chains long assumed to run Sumerian→Akkadian are being reversed: Akkadian *ikkaru* "farmer" appears to be genuinely Semitic, with Sumerian ENGAR the borrower. A 2023 study goes further, arguing that Sumerian's **formative phase** involved contact with a West Semitic facies. Watson's parallel cautions apply — *hêkal* < *é.gal* is "uncertain but feasible," and *kussû* may have travelled Semitic→Hurrian. The substrate is partly a superstrate.

**Beneath Sumerian it is murkier still.** A hypothesized pre-Sumerian substrate, "Proto-Euphratean," is credited with the river names Idiglat and Buranun, the city names Eridu, **Ur**, Larsa, Nippur and Kish, and nouns including *engar* "farmer," *simug* "smith," *nangar* "carpenter" and *damgar* "merchant" — the last of which Kramer noted had "almost universally been taken to be a Semitic hallmark." The hypothesis is disputed; Rubio's "On the Alleged Pre-Sumerian Substratum" is the standard rebuttal.



*© 2026 Noël Alejandro Bota, J.D. — Educational, historical, and genealogical research use only (Fair Use). Drafting and editing assistance with Claude-AI. Verify all claims against primary sources.*

### And which Ur?

*If* **Ur Kaśdim is the northern Ur** — Urfa, in the Harran region — the Sumerian connection disappears altogether. That identification is live precisely because a route from southern Sumerian Ur to Canaan would not naturally pass through Harran, while Genesis 11:31 has it do exactly that, and Joshua 24:2–3 places the ancestors "beyond the Euphrates."

### The one thread that survives

Not linguistic: **the moon cult.** Harran appears in **Ur III texts around 2000 BCE** as *é.ḫul.ḫul*, a temple of Sin; Ur's own patron deity was Nanna/Sin. The two cities named in Genesis 11:31 are linked by a documented cultic institution at the traditional patriarchal horizon — and the link holds whichever Ur is meant, since Harran was a major Sin centre either way.

Hurrian marks a region. Sumerian marks an era and a scribal world. Neither marks a lineage.

## Academia is not as confident about P58 as it sounds

Given the ambiguities catalogued throughout this document, the apparent confidence that J1-P58 is "the Semitic lineage" deserves scrutiny. On inspection, **that confidence lives almost entirely outside the peer-reviewed literature.**

### What the papers actually say

Read the hedges rather than the headlines:

| Source | Actual claim | Strength |
|---|---|---|
| Chiaroni et al. 2010 | J1e lineages "**might have been involved** in episodes of the expansion of pastoralists into arid habitats **coinciding with** the spread of Arabic and other Semitic-speaking populations" | Speculative; a temporal coincidence |
| Sahakyan et al. 2021 | "**Most probably**, the spread of Afro-Asiatic languages, the spread of mobile pastoralism..." | ***Hedged***; and its root-location **contradicts** Chiaroni's ***diversity result*** |
| Standard summaries | Expansion times "**overlap with**" dates for the Bronze Age spread of Semitic | ***Overlap, not causation*** |

No major paper asserts that J1-P58 is a Semitic marker. The two most-cited studies **disagree with each other** about **where P58 originated** — ***Chiaroni's highest haplotype diversity sits in the Zagros/Taurus,*** while Sahakyan's Bayesian root-location places it in Arabia, the Levant and southern Mesopotamia, explicitly contradicting the earlier diversity-based inference.



*© 2026 Noël Alejandro Bota, J.D. — Educational, historical, and genealogical research use only (Fair Use). Drafting and editing assistance with Claude-AI. Verify all claims against primary sources.*

## Evidence against, already in the literature

- J-M267 occurs in the **Canary Islands**, which are not known to have had any Semitic language [albeit, could have been brought there by external travelers].
- Most J1 outside Arabia, including in Europe, is **not** Arabic in origin despite the linguistic association
- J1-P58 is **absent from every Natufian and Neolithic Levantine male sampled**, appearing only in the Bronze Age
- In Ethiopia, J-M267 movements across the Red Sea are allegedly **older than and not limited to** the P58 form

## How the confident version propagates

Six mechanisms, none of them involving new evidence:

1. **Vivid modern frequencies.** Up to 80% in Yemen, Qatar and Sudan. Striking numbers manufacture conviction, and frequency is the one thing that is easy to measure and the least informative about origin.
2. **A seductive date coincidence.** Some J1-P58 expansion estimates land near 5,750 years ago — the same figure Kitchen et al. give for Proto-Semitic. Two independently estimated dates with wide error bars converging is suggestive, not probative.
3. **Citation drift.** "Might have been involved in episodes coinciding with" becomes "associated with" becomes "the marker for," one citation at a time, with the hedges shed at each step.
4. **The consumer-genetics channel.** Testing companies and hobbyist encyclopedias require clean labels. Hedged papers enter; branded haplogroup narratives exit.
5. **Chronological accident.** Chiaroni is 2010, **predating most Levantine ancient DNA.** *The claim entered circulation before the evidence that complicates it existed.*
6. **Demand for identity.** There is an audience for ancestral labels and essentially none for "the patrilines of the Proto-Semitic speech community are unsampled and may remain so."

## The sampling void beneath the confident claims

The regions that anchor the standard P58 narrative are the ones with almost no ancient DNA. In a compilation of 3,180 ancient Y-chromosome assignments:

| Region | Pre-Islamic ancient samples in the compilation |
|---|---|
| Saudi Arabia | **0** |
| United Arab Emirates | **0** |
| Qatar | **0** |
| Kuwait | **0** |
| Yemen | 1 — and it dates to 653 **CE** |
| Oman | 1 — 3450 BCE, and it is R1b1 |



*© 2026 Noël Alejandro Bota, J.D. — Educational, historical, and genealogical research use only (Fair Use). Drafting and editing assistance with Claude-AI. Verify all claims against primary sources.*

| Region | Pre-Islamic ancient samples in the compilation |
|---|---|
| Bahrain | 2 — both 5th–7th century CE |
| Iraq | **2 in total**, both from Nemrik 9 in the north |
| Southern Mesopotamia (Ur, Uruk, Babylon) | **0** |
| Elam / Khuzestan (Susa, Anshan) | **0** |

Sahakyan et al. root P58 in "the Arabian Peninsula, the Levant, and southern Mesopotamia." Two of those three regions are effectively unsampled for the relevant periods; the third is among the best-sampled in the world — and is duly where P58 first turns up.

This is not an argument that their model is wrong. It is an argument that the confident version of it is untestable precisely where testing would matter, and that the Levantine component of the result may be an artifact of where archaeologists have dug.

The same void constrains the opposing case. The Afro-Asiatic homeland region is in no better shape: the oldest African E in the compilation is Taforalt in **Morocco** (12,699 BCE), and the Horn and eastern Sahara produce nothing at the Proto-Afro-Asiatic horizon. Mota Cave, Ethiopia, is 2520 BCE and carries E-V38, the Niger-Congo branch.

### The asymmetry worth noticing

The ambiguities in this document — Hurrian and Sumerian layers, the contested Arpachshad etymology, two Chaldean eponyms, the northern-versus-southern Ur problem, the dating of the Table of Nations — come from philology and Assyriology, fields habituated to open questions and comfortable leaving them open.

The genetic claims reach the public through a channel that strips exactly that hedging. The underlying papers are often **more** cautious than the humanities literature; their popularizations are far less so.

### And the same test applies in reverse

This critique is not selective. The confident claim that **E-M35 is the Afro-Asiatic lineage** travels the same route from a hedged source: Ehret and Keita's observation that E1b1b's geography *coincides with* Afro-Asiatic's distribution is a correlational proposal, not a finding, and it circulates in popularized form with the same false solidity.

The correct answer to "which haplogroup is Semitic" is that the question is malformed, and that the confident answers on offer — whichever lineage they name — are artifacts of the transmission channel rather than conclusions of the research.

## Is the Table of Nations a patrilineal record?

If Genesis 10 is not a language map, the most elegant remaining hypothesis is that it tracks **patrilineal descent** — the thing Y-DNA measures. One move would then explain both mismatches: Elam under



*© 2026 Noël Alejandro Bota, J.D. — Educational, historical, and genealogical research use only (Fair Use). Drafting and editing assistance with Claude-AI. Verify all claims against primary sources.*

Shem because Elamites carried Shemite patrilines despite a non-Semitic language, Canaan under Ham despite speaking Northwest Semitic.

It is the best remaining option, and it is well motivated: Near Eastern societies were patrilineal, and the text is literally a sequence of begettings.

It fails in four places, and the first is the Table's own summary formula.

### 1. The text names its own criteria — and there are four

Genesis 10:5, 10:20 and 10:31 each close a section with the same phrase: **"according to their clans, their languages, their lands, their nations."**

Language and land are two of the four stated principles. The Table says what it is sorting by, and patrilineal descent is not the sole criterion on its own list. Standard annotation follows the text here: the relationships are based on linguistic, geographic or political grounds, citing v. 31.

### 2. Sheba and Havilah appear under both Ham and Shem

| Name | Under Ham | Under Shem |
|---|---|---|
| **Havilah** | Son of Cush (10:7) | Son of Joktan (10:29) |
| **Sheba** | Son of Raamah, grandson of Cush (10:7) | Son of Joktan (10:28) |

A patriline cannot occupy two branches. A trading region known to both spheres can. Richard Hess flags this as "a special problem," alongside a second: **the Cushites are African in verses 6–7 but Mesopotamian in verses 10–14** — the same "son" relocating a thousand miles within eight verses.

### 3. The entries are peoples, not persons

Hess observes that the Table is segmented and that names of people groups appear far more frequently than personal names: **Nimrod and the line of Shem are the only clear personal names in the chapter.**

Everything else is gentilic — Ludim, Anamim, Lehabim, Naphtuhim, Kittim, Dodanim, all Hebrew masculine plurals, plus the Jebusite, the Amorite, the Girgashite. A patriline cannot be built from plural ethnonyms. These are collectivities, grammatically marked as such.

### 4. Nimrod crosses the boundary

Genesis 10:8–12 has Nimrod, son of Cush and therefore a **Hamite**, founding Babel, Erech, Accad and Calneh in Shinar, then proceeding to Asshur and building Nineveh.

**A Hamite founds the Mesopotamian cities that are Shem's territory throughout the rest of the chapter.** Incoherent as patriliny; straightforward as a statement about Babylonian and Assyrian power.

### Genetics cannot currently test this at all

An earlier draft argued that the hypothesis fails genetically, on the grounds that Bronze Age Levantine samples are J-dominant while J1 and J2 also dominate the Zagros and Mesopotamia, so the patrilines do not separate Ham from Shem. **That argument is withdrawn — it proves nothing either way.**



*© 2026 Noël Alejandro Bota, J.D. — Educational, historical, and genealogical research use only (Fair Use). Drafting and editing assistance with Claude-AI. Verify all claims against primary sources.*

**J1 and J2 are both present across the entire region.** J1 traces to the Bronze Age in Iran at Hasanlu and Dinkha Tepe; J2 appears earlier still, with a J2b at Tepe Abdul Hosein around 8000 BCE; **J-P58 is strongly associated with ancient populations of the Caucasus and ancient Iran. *Both*** branches **occur** in the Levant, Mesopotamia, the Zagros, the plateau, Anatolia, the Caucasus and Arabia.

**J is too old and too widespread to resolve the question.** J-P209 is roughly 30,000 years old and the J1/J2 split is 20,000+ years deep. Observing that two populations "both have J" is comparable to observing that two European populations both have R — uninformative without subclade resolution. This is the identical objection raised earlier in this document against E-M35 as an Afro-Asiatic marker, and it applies here with equal force.

Testing the hypothesis properly would require dated, subclade-resolved Y-DNA from each of Shem's five territories and each of Ham's four at a comparable horizon. Susa and Anshan are unsampled; southwestern Iran is a gap the relevant authors flag themselves; the northwestern Zagros has two Middle Bronze Age individuals in total. That dataset does not exist.

**Verdict: the patrilineal reading is untestable from genetics at present — not refuted by it.** The four textual arguments above are what defeat it, and they stand independently of any genetic evidence.

**Elam is among the least-sampled major civilizations of the ancient Near East.** A 2025 study of the Iranian Plateau includes Iron Age **Neo-Elamite** samples (1100–539 BCE) from the Jubaji burial tomb in Khuzestan, plus material from Cham Papi in southwestern Ilam — but the entire study comprises 50 samples across nine sites, yielding 23 mitogenomes and 13 nuclear genomes for all of Iran across six millennia. **Susa and Anshan, the Elamite capitals, are unsampled.** The authors state the limitation directly: the lack of published ancient individuals from southwestern Iran constrains investigation of regional ancestries. A separate 2025 study of Mesopotamian aDNA reports the same gap for the adjacent northwestern Zagros, represented from the Chalcolithic to the Middle Bronze Age by only two Middle Bronze Age individuals from Hasanlu and Hajji Firuz.

The "Iranian profile" usually invoked for Elam derives from **Ganj Dareh, Hajji Firuz, Tepe Hissar and Shahr-i-Sokhta** — Zagros Neolithic, northwestern Iran, the central plateau and Sistan. None of these is Elam. Substituting a regional proxy for an unsampled population is the same error this document criticises when it is applied to haplogroups.

**What survives is the linguistic argument.** Elamite is a language isolate with no demonstrated relationship to Semitic or any other family — a conclusion drawn from texts, which Elam has in quantity from the Proto-Elamite script onward (c. 2700 BCE). The Table's placement of Elam under Shem remains a linguistic mismatch whatever future genetics shows. Note also that the recent L1-M22 study bearing on the Elamo-Dravidian question works from **modern** Y-chromosomes and states that no ancient DNA exists at the relevant time depth.

## What survives

The Table is genealogical **in form** because genealogy was the native idiom for expressing relationship in the ancient Near East. Such genealogies are segmented and fluid, adjusting to reflect contemporary political alignments rather than preserving descent records.



*© 2026 Noël Alejandro Bota, J.D. — Educational, historical, and genealogical research use only (Fair Use). Drafting and editing assistance with Claude-AI. Verify all claims against primary sources.*

The grammar is descent. The content is geography and politics. It is a relationship map written in the only notation available — which happens to look like a pedigree, and has been misread as one ever since, most consequentially in 1781.

## Iron Age Israel: more men than three, and still not Israelites

The question "what was the majority lineage in Iron Age Israel" still cannot be answered — but the record is larger than the three individuals usually cited, and the reason it fails has changed.

An earlier draft of this document put the Iron Age Levantine record at three men. A subclade-resolved compilation raises the southern Levantine count for roughly 1250–590 BCE to eleven:

| Sample | Site | Date | Y-haplogroup | Cultural attribution |
|---|---|---|---|---|
| I1934 | Tel Shadud | 1250 BCE | J1-P58 derived | Canaanite / early Israelite horizon |
| ASH066 | Ashkelon | 1234 BCE | J1-P58 derived | Philistine (Greek-affinity outlier) |
| ASH087 | Ashkelon | 1150 BCE | L2 | Philistine |
| ASH008 | Ashkelon | 1145 BCE | J1-P58 derived | Philistine |
| I4517 | Tel Megiddo | 1001 BCE | J1-P58 derived | contested Iron I–IIA |
| I2201 | Abel Beth Maacah | 939 BCE | T-CTS6280 | northern, Aramean contact zone |
| I22254 | Akhziv | 650 BCE | J1-P58 derived | Phoenician |
| I22260 | Akhziv | 650 BCE | E-M78>V22 | Phoenician |
| I22253 | Akhziv | 650 BCE | E-M78>V22 | Phoenician |
| I11788 | Akhziv | 650 BCE | E-M78 | Phoenician |
| I11794 | Akhziv | 592 BCE | E-M78 | Phoenician |

Plus the Kiryat Yearim male (**J2**), the only individual framed explicitly as Israelite or Judahite, reported in a conference paper describing "very partial results" rather than a full peer-reviewed publication.

**Eleven samples do not produce a majority, and they do not produce Israelites.** Ashkelon is Philistine. Akhziv is Phoenician. Only Megiddo, Tel Shadud and Abel Beth Maacah sit in the zone usually read as Israelite, and that reading is stratigraphic, not genetic. The securely Israelite or Judahite set remains effectively one individual.

What the larger sample does establish is **paternal heterogeneity**: P58, T, J2, L2 and at least two distinct E branches coexisting in the Iron Age southern Levant within six centuries.

### Three E lines, not one

The Iron Age Levantine E is **not** the Natufian lineage. Levantine E resolves into at least three distinct lines, on two clades, with different histories:



© 2026 Noël Alejandro Bota, J.D. — Educational, historical, and genealogical research use only (Fair Use). Drafting and editing assistance with Claude-AI. Verify all claims against primary sources.

| Clade | Levantine attestation | Character |
|---|---|---|
| **E-Z830 > M123** | Raqefet 10,750 BCE → ’Ain Ghazal 8100, 7877 BCE → Ba’ja 7025 BCE → Peqi’in 4018 BCE → Ebla 2600 BCE → Tall Sūkās 1750 BCE → Hazor 1700 BCE → Megiddo 1500 BCE (×2) | the continuous indigenous line |
| **E-M78 (V22 parent)** | ’Ain Ghazal 7615 BCE (E1b1b1a1b) | the oldest sample on the node ancestral to both V22 and V13; not the same branch as the Alalakh and Akhziv M78 below |
| **E-M78 (c1 line)** | Alalakh 1600 BCE (ALA136); Akhziv 650 and 592 BCE | a separate M78 branch; its nearest dated relatives in the compilation are East African — Kokurmatakore, Kenya, 2250 BCE (…c1a) and medieval Kulubnarti, Nubia (…c3) |
| **E-M78 > V22** | Sidon 1750 BCE; Akhziv 650 BCE (×2); Beirut 173 BCE; Chhim 619 and 623 CE; Sidon 1150 CE | Egyptian and North African affinities in its later samples — KV11, Valley of the Kings, 1170 BCE; Akhmim, Upper Egypt, 500 BCE. But the Levantine attestation comes first: Sidon 1750 BCE predates every Egyptian V22 in the compilation, and the parent node is already at PPN ’Ain Ghazal. First-sampled is not origin in either direction. |

So “E in the Levant” is at least three stories: an indigenous Z830>M123 line running Natufian to Late Bronze; a V22 branch prominent from the Middle Bronze Age onward, whose parent node is already at PPN ’Ain Ghazal and whose later samples have Egyptian affinities; and a separate M78 line at Alalakh and Akhziv whose nearest dated relatives are East African. An earlier draft lumped ’Ain Ghazal 7615 BCE with Akhziv and described a seven-millennium gap; the two sit on different M78 branches, and Alalakh (1600 BCE) carries the Akhziv branch in the interval. Any argument treating Levantine E as a single continuous presence conflates them.

Note the symmetry with the J1 case: just as J1a2b and J1a2a1b are siblings of P58 with a quite different geography, M78 and Z830 are siblings within E-M35 with quite different Levantine histories. Neither “E” nor “J1” is a unit at the resolution this argument requires.

### The elite question cannot be posed

Establishing that one lineage was elite and another common would require **status-differentiated burials sampled in numbers** — elite tombs against ordinary interments, same site, same period, enough of each to compare. No such dataset exists for the Iron Age Levant. One tomb with one male genome supports no inference about stratification.

### The Bronze Age substrate was paternally diverse

The preceding population was not a J monoculture. Tel Megiddo alone yields twenty Bronze Age males across six lineages:



*© 2026 Noël Alejandro Bota, J.D. — Educational, historical, and genealogical research use only (Fair Use). Drafting and editing assistance with Claude-AI. Verify all claims against primary sources.*

| Lineage | Count | Date range |
|---|---|---|
| J1-P58 derived | 11 | 1903–1422 BCE |
| J2 (J2a, J2b) | 4 | 2244–1401 BCE |
| E-Z830 > M123 | 2 | 1500 BCE |
| T1a | 1 | 1425 BCE |
| R1a | 1 | 1550 BCE |
| J1a3 | 1 | 1875 BCE |

Hazor adds one P58 and one E-Z830>M123, both at 1700 BCE. The northern Levant adds J2a, J2b, T1a and L1b at Sidon and Tall Sūkās. Substantial paternal diversity in the population ancestral to the Iron Age groups is itself an argument against any single lineage being "the" Canaanite or Israelite patriline.

### Two caveats the authors raise themselves

Agranat-Tamir note that both Abel Beth Maacah and Megiddo are **inland cities** whose genetic continuity across the Bronze–Iron transition "might not be representative of other sites in the region."

And "Israelite" is an **archaeological inference, not a genetic one.** Nothing in the DNA says Israelite; the attribution rests on stratigraphy and material culture. Megiddo's Iron I–IIA affiliation is contested, and Abel Beth Maacah sits in a northern zone with Aramean contact.

### The thread worth watching is no longer thin

**T in the Levant.** T was 90% of the Chalcolithic population at Peqi'in, and an earlier draft noted its reappearance at Abel Beth Maacah in the Iron Age as a single suggestive sample. It is not a single sample:

| Date | Label | Site |
|---|---|---|
| 7619 BCE | T1a | 'Ain Ghazal |
| 4384–4000 BCE | T1a1; T1a1a (×3); T1a1a1b (×3); T1a1a1b2; T1a1a1b2b1 | **Peqi'in** (9 of 10 males) |
| 1900 BCE | T1a1a1b2 | Ebla |
| 1750 BCE | T1a1a1b2b; T1a1a2a | Sidon |
| 1750 BCE | T1a2a1 | Lapithos |
| 1425 BCE | T1a1a1b2a | Tel Megiddo |
| 939 BCE | T1a1a1b2b2b1a1a2 | Abel Beth Maacah |
| 340 CE | T1a1a1b2b2b1a1a1 | Qornet ed-Deir |
| 485 CE | T1a2b | Ej-Jaouze Tombs |



*© 2026 Noël Alejandro Bota, J.D. — Educational, historical, and genealogical research use only (Fair Use). Drafting and editing assistance with Claude-AI. Verify all claims against primary sources.*

| Date | Label | Site |
|---|---|---|
| 1169 CE / 1225 CE | T1a; T1a2b2 | Sidon |

T persisted in the Levant continuously for nine thousand years, and the Abel Beth Maacah and Qornet ed-Deir individuals sit on adjacent terminal branches (…b1a1a2 and …b1a1a1).

That makes **three** long-resident Levantine lineages — E-Z830, T1a and H2 — all present before P58 arrives, and all still present after. The Levantine record is one of accumulation, not succession.

## Why J1-P58 attracts the attention it does

The earlier section on manufactured confidence omitted the largest single driver: **J1-P58 carries the Cohen Modal Haplotype.**

Skorecki et al. (1997, *Nature*) found that ***self-identified*** Cohanim share a distinctive Y-chromosome signature, interpreted as support for the tradition of patrilineal descent from Aaron. Hammer et al. (2009) refined it: an extended CMH on the **J-P58*** background predominates among both Ashkenazi and non-Ashkenazi Cohanim and is essentially absent in non-Jews, with a coalescence estimate of roughly **3,190 ± 1,090 years.**

No other Y-haplogroup is linked to a **named biblical individual and a living hereditary status**. That is an identity story nothing else in the field can match, and it has driven enormous popular and commercial interest.

### But the Cohen data does not say what the fixation implies

| Finding | Figure |
|---|---|
| Cohanim carrying J-P58* | **46.1%** — a plurality, not a majority |
| Distinct lineages among Jewish priests | **21** |
| **J2-M410* among Cohanim** | **14.4%** |
| **J2-M12 among Cohanim** | **7.4%** |
| Share covered by the top five haplogroups | 79.5% |
| **J-P58* among non-Jews surveyed** | **325 of 2,099** |

**Most** Cohanim are **not** J-P58. ***The priesthood carries 21 lineages.*** The extended CMH is specific; **P58 itself is common across the Near East** and tells you nothing on its own. And the ~3,200-year coalescence places the founder in the Iron Age — not at an Exodus horizon, and far later than Proto-Semitic.

So the P58 fixation is driven by two separate identity-laden narratives converging on one haplogroup: the Cohen/Aaron story on one side, and high modern frequencies among Arab populations on the other. Neither is evidence about Proto-Semitic, and the underlying papers claim much less than the attention suggests.



*© 2026 Noël Alejandro Bota, J.D. — Educational, historical, and genealogical research use only (Fair Use). Drafting and editing assistance with Claude-AI. Verify all claims against primary sources.*

## Four senses, routinely conflated

Only one of these has current technical standing, and it classifies languages rather than people.

| Sense | Dates from | Status | What it classifies |
|---|---|---|---|
| Genesis 10 genealogical | Iron Age text | Internally inconsistent; excludes Canaan | Peoples, **by political alignment** |
| Linguistic | 1781 | Standard; a branch of Afro-Asiatic | Languages |
| Racial | 19th century (Renan, Gobineau) | **Scientifically dead** | Purported races |
| Via "antisemitism" | Marr, 1879 | Live, but usually refers to Jews specifically – despite Arabs, Assyrians, etc. also being “Semites”. | A target of hostility |

On the fourth: "Semitism" appears in 1848 for linguistic attributes, by 1851 for "Semitic people" especially Jewish practices, and by 1870 in the specialized sense of Jewish influence in a society. Marr's *Antisemiten-Liga* (1879) fixed the compound as a euphemism for *Judenhass*. It has never meant "hostility to speakers of Afro-Asiatic languages." Arguing otherwise from the morphology is the etymological fallacy — a word's meaning is set by usage, not by its parts. A hippopotamus is not a horse.

Senses 1, 3 and 4 are claims about people. Sense 2 is a claim about grammar and lexicon. Any argument that moves between them without flagging the shift is equivocating, whoever is making it and whatever they conclude.

## Two nodes, six millennia apart

The single most common error in this debate is treating Proto-Afro-Asiatic and Proto-Semitic as the same event. They are separated by roughly 6,000–10,000 years.

**Proto-Afro-Asiatic.** No consensus on date; reconstructed range c. 16,000–10,000 BC, with Ehret dating it to 12,000–14,000 years ago. On homeland, a Northeast African origin is supported by the majority of scholars today — Ehret (Red Sea Hills / Ethiopia), Blench (southwestern Ethiopia, from Omotic diversity), Diakonoff (eastern Sahara, southern fringe). The Western Asian position is a minority view whose leading proponent is **Alexander Militarev**, who identifies Proto-Afro-Asiatic with the Levantine Natufian culture on the basis of a reconstructed farming and pastoralist lexicon.

**Proto-Semitic.** Kitchen, Ehret, Assefa & Mulligan (2009) estimate an Early Bronze Age origin approximately 5,750 years ago (c. 3750 BCE) in the Levant, with Ethiosemitic reflecting a single introduction from southern Arabia approximately 2,800 years ago. Note that Ehret — the leading African-homeland proponent — co-authored this, and the paper's tests supported an initial divergence of Akkadian from ancestral Semitic over competing hypotheses, explicitly including an African origin of Semitic.



*© 2026 Noël Alejandro Bota, J.D. — Educational, historical, and genealogical research use only (Fair Use). Drafting and editing assistance with Claude-AI. Verify all claims against primary sources.*

**A caveat that is usually dropped.** The 2009 study could not estimate when or where the ancestor of all Semitic languages diverged from Afro-Asiatic. 3750 BCE is the date of the **breakup** of Proto-Semitic into East, Central and South branches — the latest date at which it still existed as a unity. The origin of Semitic is somewhere earlier by an interval the method cannot measure.

**Semitic is also the one Afro-Asiatic branch that everyone agrees left Africa: the vast majority of Afro-Asiatic languages are indigenous to Africa, with Semitic having originated in West Asia. Under the majority homeland model, Semitic is by definition the emigrant branch.**

## The Levantine Y-chromosome record: dominance shifted, lineages persisted

The dominant paternal lineage of the Levant changed twice between the Natufian and the Bronze Age. What did **not** happen is replacement: the indigenous E lineage is attested continuously across the entire sequence.

### What shifted

| Horizon | Date | Dominant Y-haplogroup | Source |
|---|---|---|---|
| Natufian / Epipaleolithic | c. 12,000–10,750 BCE | **E** (E-Z830) | Lazaridis et al. 2016 |
| Pre-Pottery Neolithic B | c. 8300–7000 BCE | E, with T and H2 already present | Lazaridis et al. 2016 |
| Late Chalcolithic (Peqi'in) | 4500–3900 BCE | **T** — 9 of 10 males | Harney et al. 2018 |
| Bronze Age | c. 2500–1500 BCE | **J**, including P58 from 2520 BCE | Harney et al. 2018; archaeogenetic compilation |

Five Natufian specimens were analyzed for paternal lineages in the published study: one CT, one E-M35, one E-M215, one E-Z830. Harney et al. note that the Peqi'in finding contrasts with both earlier Neolithic and Epipaleolithic Levantine populations, which were dominated by haplogroup E, and later Bronze Age individuals, all of whom belonged to haplogroup J. (A subclade-resolved compilation assigns all five Raqefet Cave males to E branches; where re-called assignments differ from the published ones, the published figures are used here.)

### What persisted: E-Z830 → M123, Natufian to Late Bronze Age

An earlier draft of this document called the sequence above "two complete turnovers." **That framing was wrong and is withdrawn.** The turnovers were in *dominance*. Subclade-resolved calls show the Natufian E lineage as a single nested, progressively deepening line running the full length of the Levantine record:

| Date | Label in compilation | Branch | Site |
|---|---|---|---|
| 10,750 BCE | E1b1b1b2 ×2; E1b1b1b2a | Z830; Z830>M123 | Raqefet Cave (Natufian) |



© 2026 Noël Alejandro Bota, J.D. — Educational, historical, and genealogical research use only (Fair Use). Drafting and editing assistance with Claude-AI. Verify all claims against primary sources.

| Date | Label in compilation | Branch | Site |
|---|---|---|---|
| 8100 BCE | E1b1b1b2 | Z830 | ’Ain Ghazal |
| 7877 BCE | E1b1b1b2a1a | M123>M34 | ’Ain Ghazal |
| 7025 BCE | E1b1b1b2a1a1 | M136 | Ba’ja |
| **4018 BCE** | E1b1b1b2a1a1 | M136 | **Peqi’in** |
| **2600 BCE** | E1b1b1b2a1a1a1a1f1a | deep M123 line | **Ebla* [predates J1-P58’s 2520BCE]** |
| 1750 BCE | E1b1b1b2a1a1a; …a1a1a2 | deep M123 line | Tall Sūkās |
| **1700 BCE** | E1b1b1b2a1a1a1a1**f1** | deep M123 line | **Hazor** |
| **1500 BCE** (×2) | E1b1b1b2a1a1a1a1**f1** | deep M123 line | **Tel Megiddo** |

The Hazor and Megiddo individuals share the f1 node with the Ebla individual’s f1a. That is direct Levantine paternal continuity across roughly nine millennia — through the Proto-Semitic horizon and into the period of attested Semitic writing. E-M35 is present even at Peqi’in, the site that is 90% T.

**But the f1 node is not Levant-exclusive.** Its earliest member in the compilation is ART015 at Arslantepe in eastern Anatolia (3262 BCE, f1a), some 660 years before Ebla. Oylum Höyük adds two more on the same node (I14784, 1745 BCE; I14782, 1725 BCE), and the line travels west with Phoenician settlement — Kerkouane, Tunisia, 550 BCE (f1) and Carthage 325 BCE (f1b1a). The wider M123 line is also on the upper Tigris at Tatika (3005 BCE), in Bronze and Iron Age Armenia on the sibling e1 node, and at Gonur in Turkmenistan (1931 BCE). Continuity in the Levant is real; exclusivity to the Levant is not, and the f1 node is first sampled in the corridor rather than on the coast. By this document’s own rule that does not make eastern Anatolia its origin — but it does stop the f1 continuity being read as a purely Levantine story.

**T and H are not Chalcolithic adopters either.** H2 is attested at ’Ain Ghazal by 8100 BCE and T1a by 7619 BCE — some 2,600 years after the Natufian E samples, but roughly five millennia before Peqi’in. All three lineages are long-resident in the Levant; what changes over time is their relative frequency, not their presence.

**Autosomal ancestry moved on a different schedule from the Y-chromosome.** The Peqi’in population models as ~57% Levant Neolithic, ~26% Anatolian Neolithic, ~17% Iran Chalcolithic — majority Natufian-descended ancestry carried by men who were 90% non-Natufian in patriline. Agranat-Tamir et al. found Bronze Age “Canaanite” groups descending from local Neolithic populations plus populations related to the Chalcolithic Zagros or Bronze Age Caucasus, with people related to these individuals contributing to all present-day Levantine populations.

**J1-P58 is a demonstrable latecomer, and its floor is now precise.** No J1 of any kind appears in a Levantine sample before **ETM012 at Ebla, 2520 BCE *(2574–2467 calBCE). P58 was absent from all Natufian and Neolithic Levantine males examined, emerging in the Bronze Age alongside Chalcolithic-Iranian-related ancestry. That associated ancestry penetrated the Levant between 3,300 and 5,900 YBP and Arabia between 2,000 and 3,500 YBP.***



*© 2026 Noël Alejandro Bota, J.D. — Educational, historical, and genealogical research use only (Fair Use). Drafting and editing assistance with Claude-AI. Verify all claims against primary sources.*

**The Proto-Semitic horizon itself remains unsampled.** The mainland Levantine sequence runs Peqi'in (c. 4000 BCE) → Ebla (2600 BCE) with nothing in between. Kitchen et al.'s 3750 BCE estimate falls squarely in that gap, bracketed by a T-dominant population before it and an E-plus-J population after it.

**One caveat on Peqi'in.** It is one cave, 22 individuals, ten males, one region, one material culture. The paper also notes the Peqi'in population contributed differently to later Bronze Age groups, one of which cannot plausibly have descended from the same population — so it may be a partly dead-end regional population rather than the direct forebear of Bronze Age Levantines.

## The thesis: E-M35 as founder lineage

The position under examination runs roughly as follows.

1. Afro-Asiatic originated in Northeast Africa (the majority view).
2. E-M35 is of ultimate African origin and is frequent across Afro-Asiatic-speaking populations.
3. The Natufians carried E-Z830 and represent the Levantine edge of an African Afro-Asiatic expansion.
4. Semitic descends from that Afro-Asiatic trunk; it cannot have been spoken by anyone who did not receive it from the trunk's carriers.
5. Therefore E-M35 is the founder lineage, and T1a and H2a were adopters.
6. J1-P58, which dominates Semitic-speaking populations today, is a secondary Bronze Age arrival from the Zagros and has no claim to founder status.

Points 1, 2, 3 and 6 are well supported. Point 6 in particular is better grounded than the popular account, which tends to read modern J1 frequency backwards into antiquity.

**Point 5 needs amending on the evidence.** T1a and H2 are not adopters in any meaningful sense: H2 is attested in the Levant at 'Ain Ghazal by 8100 BCE and T1a by 7619 BCE, millennia before the Chalcolithic. E is attested earlier still, in the Natufian, so the *ordering* in the thesis holds — but the three lineages are long co-resident, not a founder followed by latecomers.

The argument fails at points 4 and 5, and the failure is structural rather than evidential. Those two steps require that a language family's affiliation can be read off a Y-haplogroup — the inference pattern the rest of this document tests.

## The case for, at its strongest

Stated fairly, the thesis has real support and should not be dismissed. Three of the strongest arguments come from subclade-resolved ancient DNA rather than from modern frequencies.

**E-M35 is present at the first Semitic-literate city, and it is the earlier of the two samples there.** Ebla produced Eblaite, one of the two oldest attested Semitic languages. In the compilation, ETM010 at **2600 BCE** carries the deep Levantine E-Z830>M123 line (E1b1b1b2a1a1a1a1f1a); ETM012, the earliest P58 anywhere in the dataset, is **2520 BCE**. At the one site where Semitic speech is documented rather than inferred, E-M35 is attested, and in this sample sequence it precedes P58 by roughly eighty years. This is not a frequency correlation — it is physical presence at the relevant place and time.



*© 2026 Noël Alejandro Bota, J.D. — Educational, historical, and genealogical research use only (Fair Use). Drafting and editing assistance with Claude-AI. Verify all claims against primary sources.*

**One E clade is attested across three Afro-Asiatic branches.** E-Z830 (E1b1b1b2) splits into sub-branches that sit in three distinct Afro-Asiatic-speaking zones, each with ancient attestation:

| Z830 sub-branch | Region | Afro-Asiatic branch | Ancient attestation |
|---|---|---|---|
| E1b1b1b2a = **M123** | Levant | Semitic | Raqefet 10,750 BCE → Megiddo 1500 BCE |
| E1b1b1b2b = **M293** | East Africa | Cushitic | Kokurmatakore 2250 BCE; Lukenya Hill 1549 BCE; Kisima Farm 1017 BCE; Naishi 650 BCE; Rigo 643/502 BCE (all Kenya) |
| Z830 line | Egypt | Egyptian | Nuwayrat, Old Kingdom, 2700 BCE *(call needs verification — M293 in Egypt is unexpected)* |

This is a materially better argument than Ehret and Keita's modern-frequency observation. A single clade with dated sub-branches in Semitic, Cushitic and Egyptian territory is the shape the evidence would take if E-Z830 had travelled with early Afro-Asiatic. It still does not prove that it did — contemporaneous presence is not transmission — but it is the strongest form of the case, and it should be stated as such.

**One qualification the compilation imposes.** M123 is not confined to Semitic-speaking territory. Beyond the Levant it appears at Arslantepe and Tatika in Anatolia, among seven Bronze and Iron Age Armenians in the Hurro-Urartian zone, at Gonur in the Bactria–Margiana complex, and in nine Iron Age males from the Swat Valley. The "Semitic" label on the Levant row describes where those samples were found, not a property of the branch — the same caution this document applies to P58.

**Levantine E continuity is real, not residual.** The Natufian lineage runs unbroken to the Late Bronze Age, including at Peqi'in itself, with Hazor and Megiddo sharing a sub-branch node with Ebla. Any account of Levantine population history that treats E as having been displaced is wrong on the evidence.

**The Natufian E is genuinely suggestive.** Lazaridis et al. themselves noted that craniometric analyses have suggested Natufian migration from north or sub-Saharan Africa, a result finding some support from Y-chromosome analysis showing haplogroup E, of likely ultimate African origin, which had not been detected in other ancient West Eurasian males. The authors raised the African connection; it is not an outside imposition on the data.

**The anti-presentist critique is correct and well-applied.** Modern frequency does not locate ancient origin. Chiaroni et al. found the highest J1e haplotype diversity in the Zagros/Taurus region while J1e *frequency* increases toward the peripheral Arabian Peninsula — a southerly pattern of decreasing expansion-time estimates consistent with serial drift and founder effects. High frequency with low diversity is the signature of peripheral founder effect, not of ancestral centrality. The Bedouin J1 star-burst is drift, not origin.

**The critique of Bayesian root-location is reasonable.** Sahakyan et al. (2021) place J1a1a1-P58's root in the Arabian Peninsula, the Levant and southern Mesopotamia, explicitly contradicting earlier Y-STR



*© 2026 Noël Alejandro Bota, J.D. — Educational, historical, and genealogical research use only (Fair Use). Drafting and editing assistance with Claude-AI. Verify all claims against primary sources.*

diversity results. But that is an inference from modern whole-Y sequences, and it sits badly against ancient DNA showing P58 absent from the Levant until the Bronze Age while E, T and H are all attested there earlier. Modern sequence distributions are exactly what aDNA keeps overturning.

**The "there were already lineages there to speak it" point is sound.** Whoever spoke Proto-Semitic in the Levant at 3750 BCE was drawn from E, T or H — not from J, which had not arrived.

## Testing the thesis: a mixed scorecard

Each of these can be checked against published data. The scorecard is no longer uniform: two tests now go to the thesis, three go against it, and one cannot be run.

| Test | Prediction if E carried Afro-Asiatic into the Levant | Result |
|---|---|---|
| Natufian autosomal ancestry | African component present | **Against** — no affinity to sub-Saharan Africans; all tests non-significant, $\|Z\| < 3$ |
| Levantine patriline at the Proto-Semitic horizon | E-dominant | **Mostly against** — 90% T at Peqi'in (4500–3900 BCE), though E-M136 is present (1 of 10) |
| Patriline at the first site with attested Semitic writing | E present | **For** — at Ebla, E-Z830>M123 is attested at 2600 BCE, preceding the earliest P58 (2520 BCE) at the same site |
| A shared E clade across Afro-Asiatic branches | Yes | **For** — E-Z830 sub-branches in the Semitic Levant (M123), Cushitic East Africa (M293) and Egypt, all ancient |
| Afro-Asiatic branches uniformly E-dominant | Yes | **Against** — Chadic is up to 95.5% R1b-V88 |
| Y-DNA from the Afro-Asiatic homeland population | Available and E-dominant | **Cannot be run** — never sampled; no data exists |

Three further problems are chronological or structural rather than empirical.

**E-M35 predates Afro-Asiatic.** Its coalescence age is ~23,900 BP, with a possible origin near 34,800 BP. Proto-Afro-Asiatic is at most c. 16,000–10,000 BC. E-M35 was already old, diversified and distributed across North Africa and the Near East before the language family existed. The compilation sharpens this: the Iberomaurusians of **Taforalt, Morocco, carry E-M78 at 12,699–12,271 BCE**, contemporary with the E-Z830 Natufians. E-M35's major branches were already split and geographically separated at the Proto-Afro-Asiatic horizon. A lineage that predates a family and is already differentiated across its future range cannot by itself be that family's tracer.



*© 2026 Noël Alejandro Bota, J.D. — Educational, historical, and genealogical research use only (Fair Use). Drafting and editing assistance with Claude-AI. Verify all claims against primary sources.*

**E is not Afro-Asiatic-specific.** E1b1a (E-V38) is the Niger-Congo/Bantu lineage, carried by hundreds of millions of speakers of an unrelated family — and it is attested anciently, at Mota Cave, Ethiopia, 2520 BCE. E1b1b reaches ~10% in some Khoisan groups. By population, E1b is overwhelmingly a non-Afro-Asiatic marker.

**The direction of travel is contested.** The Moroccan Early Neolithic individuals from Ifri n'Amr ou Moussa carried E-M81 (E1b1b1b1a, under L19) — the Berber branch, not the Levantine Z830 line — and the majority of their genomes traced back to Natufian and PPN ancestry, indicating a recent intrusion into North Africa **from** the Levant.

**The arithmetic objection.** Even granting the thesis entirely: descent from a population 5,750+ years ago spans roughly 230 generations. Within a regionally connected population, nearly everyone who left descendants at all is an ancestor of nearly everyone downstream. "Descended from the Proto-Semitic speech community" does not partition anyone from anyone. Y-haplogroups are the one component of ancestry that does *not* fan out, which is precisely why building a criterion on them manufactures a sharpness the actual descent structure lacks.

## Where the argument moved

Both sides gave ground on specifics. Recording this matters, because an argument that never moves is not being tested.

**Conceded against the skeptical position:**

- The claim that the **Levant at 3750 BCE** already carried J was **wrong**. J is unattested in the Levant before the Bronze Age. **The correct list at that horizon is E, T, H.**
- Diakonoff was misattributed to the Western Asian homeland camp. He held an **African** origin for Afro-Asiatic (eastern Sahara / northeast Africa), revising only the Common Semitic homeland toward the Delta and Sinai. Militarev is effectively the sole leading Levantine proponent.
- The homeland question is not evenly split. **Northeast Africa is the majority view.**
- The Chadic counterexample was presented as settled when the R-V88/Chadic association is **actively disputed** (see next section).
- An African-homeland-then-Levantine-expansion model does not require Militarev, and the "you need both homelands" objection was aimed at a position not being held.

**Conceded against the E-M35 thesis:**

- Excluding J on ancient-DNA grounds sharpens the picture to one where **T is dominant** at the Proto-Semitic horizon. The same evidence cannot exclude J and exempt T.
- The anti-presentist rule used to demote J1-P58 demotes E-M35 by the identical step when applied evenly.
- **3750 BCE** is a **breakup** date, not an origin date, so "the first speakers" is an undated node in an unknown location.
- There are no "first speakers" to point at: the first Semitic speakers and the last speakers of pre-Semitic Afro-Asiatic are the same people. The boundary is a retrospective classificatory decision.

**Conceded after the subclade-resolved compilation:**



*© 2026 Noël Alejandro Bota, J.D. — Educational, historical, and genealogical research use only (Fair Use). Drafting and editing assistance with Claude-AI. Verify all claims against primary sources.*

- The "two complete turnovers" description of the Levantine Y-record was **wrong**. E-Z830>M123 runs continuously from the Natufian to Late Bronze Age Megiddo, and is present at Peqi'in itself. The turnovers were in dominance, not presence.
- E-M35 **is** attested at Ebla, the first Semitic-literate city, and in that sample sequence it precedes the earliest P58 by eighty years. This is the strongest single datum the thesis has, and it is not a frequency argument.
- A single E clade — Z830 — has dated sub-branches in the Semitic Levant (M123), Cushitic East Africa (M293) and Egypt. That is a materially better argument for an Afro-Asiatic association than the modern-frequency correlation this document criticises.
- The objection that P58's earliest attestation is the northern Levant rather than the Zagros was a **first-sampled-equals-origin error** and is withdrawn. The upstream J1a2a node is at Tell Kurdu by 5652 BCE with non-P58 descendants in Turkmenistan, Iran and Anatolia; the sibling J1a2b branch is Caucasus-centred; the deepest J1 is at Satsurblia (11,332 BCE) and Nemrik 9 (8750 BCE); and the Chalcolithic-Iranian autosomal signal arrives with P58. The clade is rooted in the Zagros–Caucasus–Anatolian arc.
- Arabia, southern Mesopotamia and Elam are **effectively unsampled**, which makes the confident version of any P58 origin story — Arabian, Levantine or Zagros — untestable in two of the three regions that matter.
- The **sibling-branch geography** was then over-weighted in the other direction. J1a2a1b is J-ZS6599, which parts from the P58 line at J-PF4832 around 12,000 BCE — nine millennia of separate history before either branch is sampled. Presenting its Turkmenistan and Iranian attestations as direct corroboration of P58's rooting repeated the document's own cardinal error, and is withdrawn. The autosomal Chalcolithic-Iranian co-arrival carries that argument; the uniparental sibling data does not.

## The Chadic case: the premise fails under either model

Chadic is the cleanest worked example in the whole debate, and it is worth laying out because it is genuinely contested.

**The frequencies.** R1b-V88 reaches 95.5% among the Ouldeme (N. Cameroon), 82.4% among the Mada, 75% among other Chadic groups, and 20% among the Hausa. R1b is a back-migration from Asia.

**There is no ancient DNA for R-V88 in Africa.** Cruciani et al. inferred a mid-Holocene trans-Saharan migration of proto-Chadic speakers from an R-V88 coalescence of 9.2–5.6 kya; Haber et al. estimated the Chad V88 chromosomes coalesced 5,700–7,300 years ago. Both are coalescence estimates from modern samples. Coalescence is not arrival.

**The aDNA that exists is European.** Marcus et al. detected R1b-V88 markers in 11 of 30 ancient Sardinian males from the Middle Neolithic to the Nuragic, plus 10 mainland European ancient samples. The subclade marker R1b-V2197 — found today in Sardinians and most African R1b-V88 carriers — is derived only in the Els Trocs individual (Pyrenees, c. 5100 BCE) and two ancient Sardinians. Basal V88 appears in the Mesolithic Balkans and Neolithic Europe; the authors conclude V88 traces back to eastern Europe.



*© 2026 Noël Alejandro Bota, J.D. — Educational, historical, and genealogical research use only (Fair Use). Drafting and editing assistance with Claude-AI. Verify all claims against primary sources.*

**The Chadic link is disputed.** Kulichová et al. found V88 is not specific to Chadic speakers, occurring widely in the Sahel and especially among Arabic speakers. Shriner & Rotimi proposed it entered via Baggarization, spreading from the 13th century onwards and reaching Chadic populations such as the Hausa and Mada as recently as ~300 years ago. The Chad genetic-history study concludes that Baggara Arabs spoke Arabic, not Chadic, implying people carrying R1b-V88 were not responsible for the spread of Chadic languages, which may have spread approximately 3,700 years ago.

**Why it breaks the premise either way:**

- **Cruciani's model:** a Eurasian back-migrant lineage carried an Afro-Asiatic branch → E is not the family's marker.
- **Shriner & Rotimi's model:** a foreign lineage replaced the patriline to 95% within a few centuries, carried by Arabic speakers, while the population went on speaking Chadic → near-total paternal replacement with zero language change.

The dispute removes Chadic as evidence about Afro-Asiatic *origins*. **It strengthens Chadic as evidence that a population's dominant Y-haplogroup can be almost wholly replaced without touching what they speak** — the same decoupling seen at Peqi'in, on a different continent in a different millennium.

## Shem and Semitic were never the same category

The apparent deadlock dissolves once the biblical category and the linguistic one are pulled apart. They point in opposite directions, and both readings can be correct at once.

### Shem's territory is eastern

| Shem's son | Location | Direction from Canaan |
|---|---|---|
| Elam | Susiana, southwestern Iran — the **Zagros** | East |
| Asshur | Assyria, upper Tigris | Northeast |
| Arpachshad | Mesopotamian; sometimes linked to Arrapḫa (Kirkuk region), though the etymology is contested. The line leading to Abraham and Ur Kaśdim | East |
| Aram | Syria and Aram-Naharaim, upper Mesopotamia | Northeast |
| Lud | Uncertain — Lydia, or Lubdu in Assyria | North / Northeast |

Genesis 10 makes Shem the **eastern bloc**: Mesopotamia, the Zagros piedmont, upper Syria. The Mediterranean coast (Canaan), Egypt (Mizraim) and Cush all go to Ham. The text places the Shemites exactly where Iranian-Chalcolithic-related ancestry and J1-P58 demonstrably originated, and places the Levant — where Proto-Semitic was actually spoken — under Ham.

Abraham's itinerary reinforces it. Ur, then Harran, in Aram-Naharaim. Deuteronomy 26:5 confesses a wandering Aramean father; Genesis 31:47 has Laban's branch naming things in Aramaic. The patriarchal line is depicted as eastern in origin, Aramaean in speech, and arriving in Canaan from outside.



*© 2026 Noël Alejandro Bota, J.D. — Educational, historical, and genealogical research use only (Fair Use). Drafting and editing assistance with Claude-AI. Verify all claims against primary sources.*

## So the convergence is real

The textual geography of Shem and the arrival path of J1-P58 point to the same region. On the Bible's own logic, **J1-P58 fits Shem better than E-M35 does.**

**A withdrawn objection.** An earlier draft argued against this convergence on the ground that P58's earliest *attestation* is not the Zagros but the northern Levant and southeastern Anatolia — ETM012 at Ebla (2520 BCE) and a near-contemporaneous sample at Oylum Höyük in Kilis province (2500 BCE). The observation is correct; the inference drawn from it was not, and it is withdrawn. First-sampled is not origin, and it fails for exactly the reason this document gives elsewhere for why coalescence is not arrival. Four lines of evidence root the clade north and east:

**The phylogeny around P58 runs north and east.** Every branch of J1 other than P58 itself is attested first in the Caucasus, northern Mesopotamia, Anatolia or Central Asia:

| Node | SNP equivalent | Divergence from the P58 line | Ancient attestation | Regions |
|---|---|---|---|---|
| J1b | — | at or near the J1 root | **Satsurblia 11,332 BCE** | Caucasus (Georgia) |
| J1 basal | unresolved | branch unknown | **Nemrik 9, 8750 BCE** | N. Mesopotamia (Iraq) |
| J1a2a basal | upstream of PF4832 | ancestral, not divergent | **Tell Kurdu 5652 BCE**; Geoksyur 3075 BCE; Titriş Höyük 2150 BCE; Liarsang-Bon 150 BCE–0 CE | Anatolia; Turkmenistan; Iran |
| J1a2a1b → …b1a | **J-ZS6599** | at **J-PF4832, c. 12,000 BCE** | Geoksyur 3200, 3100, 3075, 3013 BCE; Gonur 2110 BCE; Hasanlu 1354 BCE; Aksu-Ayuly 671 BCE; Mayemer II 550 BCE; Kuriki Höyük 451 BCE | Turkmenistan; Iran; Kazakhstan; Anatolia |
| J1a2b + subclades | — | under J1a2 | Khvalynsk 5031 BCE; Arslantepe 3310 BCE; Taribana 3300/3294 BCE; Shah Tepe 3150 BCE; Balichi-Dzedzvebi 3000 BCE; Velikent 2900 BCE; Dzhoghaz 2459–2375 BCE; Ginchi 2320/2169 BCE | lower Volga; Anatolia; Caucasus; Iran |
| J1a2a1a2 | **J-P58** | — (the branch in question) | **Ebla 2520 BCE**; Oylum Höyük 2500 BCE | N. Levant; SE Anatolia |



*© 2026 Noël Alejandro Bota, J.D. — Educational, historical, and genealogical research use only (Fair Use). Drafting and editing assistance with Claude-AI. Verify all claims against primary sources.*

## How much weight the sibling geography can bear: less than it looks

**The nested labels create an illusion of closeness.** J1a2a1a2 and J1a2a1b look like near cousins. They are not. The compilation's J1a2a1b corresponds to **J-ZS6599**, whose last common ancestor with P58 is **J-PF4832**, dated to roughly 12,000 BCE — TheYtree gives ZS6599's formation at about 12,900 years ago and YFull about 13,100. The P58 side of that node then runs PF4832 → L136 → Z1874 → P58, with P58's own most recent common ancestor around 7,750 BCE.

So the two branches had some **nine millennia of independent history** before either is observed. A J1a2a1b1a individual at Geoksyur in 3200 BCE tells us where *that* branch was in 3200 BCE. It does not constrain where the P58 line sat during the intervening 8,800 years.

This is the same objection this document raises elsewhere — that coalescence is not arrival, and that a clade too old and too widespread cannot resolve a question of this granularity — and it applies to the eastern-rooting argument as surely as to anything else. An earlier version of this section presented the sibling geography as straightforward corroboration. It is not, and that presentation is withdrawn.

**What survives the objection, in order of strength:**

1. **The autosomal co-arrival.** P58 enters the Levantine record together with Chalcolithic-Iranian-related ancestry. This is a genome-wide signal, not a single-marker inference, and it is therefore immune to the deep-divergence problem above. It is the strongest leg of the eastern case.
2. **An ancestral node in Anatolia.** The basal J1a2a call at Tell Kurdu (5652 BCE) sits upstream of the PF4832 split, so it is genuinely ancestral to P58 rather than collateral — though a basal call on an early sample may simply reflect limited coverage.
3. **The pattern across branches.** Every observed J1 branch other than P58 is first attested in the Caucasus, northern Mesopotamia, Anatolia or Central Asia. Each individual datum is weakened by its divergence date, but the consistency of the pattern is itself an argument, and nothing in the record points P58's ancestry south or west.

*Not verified here:* the precise tree position of **L620**, reported as upstream and distributed across Turkey and Turkmenistan. The claim is consistent with the topology above but was not independently confirmed for this document.

**And the autosomal signal corroborates it.** P58's appearance in the Levant coincides with the influx of Chalcolithic-Iranian-related ancestry — a statement about genome-wide ancestry, not about one marker, and therefore not vulnerable to the single-lineage objections this document raises elsewhere.

So the clade is rooted in the **Zagros–Caucasus–Anatolian arc**, and P58 is the one branch that turns southwest. Ebla and Oylum Höyük are where it is first *caught*, not where it began. Both sit in upper Syria — which the Table of Nations assigns to **Aram**, still Shem's bloc. The convergence survives; what moves is the quadrant, from the Elamite southeast to the Aramaean northwest.

## But the Table of Nations postdates the expansion

This is the caveat that keeps the convergence from being evidence of origin.



*© 2026 Noël Alejandro Bota, J.D. — Educational, historical, and genealogical research use only (Fair Use). Drafting and editing assistance with Claude-AI. Verify all claims against primary sources.*

Genesis 10 is a first-millennium BCE composition — the "Ur Kaśdim" anachronism alone (Chaldeans are attested from the 9th century BCE) places its horizon in the Neo-Assyrian or Neo-Babylonian period at the earliest. By then the J1-P58 expansion and the Zagros-related gene flow into the Levant had been complete for well over a millennium.

An author writing in the 6th century BCE, sorting the peoples of his world into an eastern bloc and a coastal bloc, would produce exactly this map — because that is what the demography already looked like **after** the expansion. The text plausibly records the outcome, not a memory of the origin.

That is sixth-century presentism: reading present distribution as ancestral distribution. It is the same error identified elsewhere in this document in Chiaroni's diversity argument and Sahakyan's root-location model. The Table of Nations is simply the oldest surviving instance of it.

### Abraham's language: three candidates

- **Sumerian** — very unlikely. Already dead as a spoken language by roughly 1800 BCE, surviving as a liturgical and scribal register. And "Ur of the Chaldees" is anachronistic regardless.
- **Elamite** — the text puts Elam under Shem, but Elamite is a language isolate, unrelated to Semitic or to Afro-Asiatic. This is one of the places the Table fails outright as a linguistic taxonomy. Possible.
- **Aramaic** — what the text itself points toward, and the only candidate with internal support. But Aramaic is attested as a distinct language only from roughly the 10th–9th centuries BCE, making this too a first-millennium projection onto a second-millennium setting.

Open to speculation. The text gives its authors' guess, not data.

### The resolution

Two categories, pointing opposite ways:

| | Orientation | Fits J1-P58? | Fits E-M35? |
|---|---|---|---|
| **Shem** (Genesis 10) | **Eastern Origin: Mesopotamia, Zagros, migration to upper Syria** | **Yes** | **Not as Strongly**<br>**(Counter-argument: E and T1a share Shem's territory with P58 — see the note below)** |
| **Semitic** (linguistic family, c. 3750 BCE) | **Levantine** | **No — absent until the Bronze Age** | **Partially — present, but not dominant at that horizon** |



*© 2026 Noël Alejandro Bota, J.D. — Educational, historical, and genealogical research use only (Fair Use). Drafting and editing assistance with Claude-AI. Verify all claims against primary sources.*

## The E1b1b counter-argument, restated

The objection in the table is right in spirit, but its premise needs correcting. **T1a is attested in ancient Iran; E1b1b is not — yet.** The compilation holds 42 ancient Iranian males, from Belt and Hotu Caves through Tepe Abdul Hosein, Tepe Guran, Hajji Firuz, Seh Gabi, Shah Tepe, Tepe Hissar, Shahr-i Sokhta, Hasanlu and Dinkha Tepe. None is E. Three are T1a: sha014 at Shah Tepe on the Gorgan Plain (3150 BCE, T1a3b2) and two at Tepe Hissar (I2512, 2888 BCE, T1a3b; I2514, 2397 BCE, T1a3b2).

**The E sits all around the plateau, just not on it.** The M123 line appears at Arslantepe (3262 BCE), on the upper Tigris at Tatika (3005 BCE), among seven Armenian males from 1527 to 250 BCE, at Gonur in Turkmenistan (1931 BCE) and in nine Iron Age males from the Swat Valley (1266–861 BCE); a Z830 male is at Farkhor in Tajikistan by 2500 BCE. The Iranian plateau is a hole in the middle of that ring. And at Gonur — the Bactria–Margiana complex noted earlier in connection with Persian and Elamite contact — E and T1a do turn up together: I2085 (E-M123 line, 1931 BCE) and I1781 (T1a3b, 1882 BCE). That is the nearest place to Iran where the pairing is attested.

**Inside Shem's territory, E sits beside T1a in the upper Mesopotamian corridor.** The corridor table later in this document records the Ebla line of E-M123 alongside T1a and P58 at Ebla (2600–1850 BCE) and Oylum Höyük (2500–1633 BCE), and an E-M78 alongside T1a and P58 at Alalakh (1911–1428 BCE). Upper Syria is Aram in Genesis 10. Inside Shem's own bloc, E, T1a and P58 were co-resident in the same cities in the same centuries.

**What that does to singular "Shem" narratives.** Any account that equates Shem with one patriline — P58 or E-M35 — has to explain why the territory Genesis assigns to Shem was paternally plural at every sampled horizon, and why E and T1a — on other branches — also flank the Elamite and Persian east. The "Yes" in the table measures fit with the arrival path of P58. It does not make P58 the only lineage in Shem's geography, and the evidence does not let it.

**Two cautions.** First, the absence of E in Iran is a sampling result, not a finding: Susa, Anshan and southwestern Iran are unsampled (see the Elam discussion above). The ring makes a plateau E more likely; it does not establish one. Second, the Iranian T1a is the wrong T1a for a Levantine link. All three Iranian samples are T1a3, a branch whose other members in the compilation are at Çatalhöyük (6550 BCE), Pietrele in Romania, Ipatovo in the north Caucasus, Bronze Age Crete and Gonur — while the Levantine series at Peqi'in, Ebla and Megiddo is T1a1. Iranian T1a ties the plateau to Anatolia and the steppe edge, not to the Levant. Likewise the Armenian M123 sits on the e1 node, a sibling of the Levantine f1 rather than a member of it.

Schlözer welded these together in 1781 by borrowing Shem's name for a language family. They have been discussed as one thing ever since, and almost every dispute about "who is a Semite" turns out to be an argument in which one party means Shem and the other means Semitic.

*J1-P58 can be Shemite in the Genesis sense **and** a latecomer with respect to Proto-Semitic.* **Both** are true simultaneously, because the two words were joined by an eighteenth-century filing decision rather than by anything in the world.



© 2026 Noël Alejandro Bota, J.D. — Educational, historical, and genealogical research use only (Fair Use). Drafting and editing assistance with Claude-AI. Verify all claims against primary sources.

### Standing caveat

Everything above is provisional. Ancient DNA is a fast-moving field and the Levant, the Zagros and especially Northeast Africa remain thinly sampled at the horizons that matter most. A single well-dated burial from the Afro-Asiatic homeland region, or a Chalcolithic Levantine site outside Peqi'in, could revise any of this — including the sequence in the table above. The Chadic section of this document is one live demonstration: a finding published in 2010 and widely repeated was substantially complicated by work appearing over the following decade. The revisions recorded in **Where the argument moved** are another — a single subclade-resolved dataset overturned this document's own description of the Levantine Y-record and withdrew one of its objections outright.

These distinctions — descent versus speech, ***Shem versus Semitic***, coalescence versus arrival, frequency versus origin — are routinely glazed over in both popular and partisan treatments of this subject. They emerged here only from working through the primary literature case by case, and they are the part of this document most likely to survive the next round of data.

## What would settle it

The disagreement is currently unfalsifiable rather than refuted, which is a different thing and worth stating plainly.

**Would support the E-M35 thesis:**

- Ancient DNA from a Northeast African population at the Afro-Asiatic horizon (c. 12,000–10,000 BCE) showing E-M35 dominance. This is the decisive test and the data does not exist — the eastern Sahara and Horn at that period have almost no recoverable aDNA, and preservation conditions are poor.
- Levantine samples at or near 3750 BCE showing E rather than T, establishing Peqi'in as a regional outlier rather than representative.
- E-dominant samples from Bronze Age sites where Semitic is epigraphically attested. E presence at such a site is now established (Ebla, 2600 BCE); dominance is not, and Ebla's own sequence has P58 arriving eighty years later and J1 outnumbering E across the site's later phases.

**Would further undermine it:**

- More Chalcolithic Levantine sites confirming T-dominance across the region.
- African aDNA at the Afro-Asiatic horizon showing lineage diversity rather than E-dominance.

**Would not settle anything either way:** modern frequency distributions, Bayesian root-location from contemporary sequences, or Y-STR coalescence estimates. All three have now been overturned by ancient DNA at least once in this specific debate.

**The standing open question.** No population identifiable as Proto-Afro-Asiatic speakers has ever been sampled. The trunk's patrilines are simply not in evidence, and an unobserved node cannot anchor a definition of who anyone is.

**And the framing question survives all of it.** Even with perfect data, ***"Semite" remains a label coined in 1781 from a genealogy that classifies Hebrew as Hamitic.*** The historical questions — who the Natufians



© 2026 Noël Alejandro Bota, J.D. — Educational, historical, and genealogical research use only (Fair Use). Drafting and editing assistance with Claude-AI. Verify all claims against primary sources.

were, what the Chalcolithic turnover meant, when and whence J1-P58 arrived — are real and answerable. The step from any of those answers to entitlement to a name is the one that keeps failing, and it fails on the arithmetic of descent before it ever reaches politics.

## The upper Mesopotamian corridor: the question this document defers

This document has brushed repeatedly against a zone it has not examined on its own terms: the upper Tigris, the Khabur and the Anatolian–Zagros piedmont, where Semitic, Hurrian, Indo-Anatolian and Iranian-related ancestry all overlap. The Hurrian sections above concluded that the Hurrian material is contact rather than descent, and that verdict stands. But contact is not a null finding when the contact zone is also the corridor through which every lineage discussed here reached the Levant.

Two results from the subclade-resolved compilation make the corridor worth separate treatment.

**Alalakh is the type site for lineage co-residence.** Across the Tell Atchana entries, roughly twenty males span six haplogroups in a single Mitanni-period city — and the surrounding sites show the same mixing at different dates:

| Site | n | Date range | Composition |
|---|---|---|---|
| Arslantepe (E. Anatolia) | 13 | 3817–2392 BCE | **J2a ×8**, H2b ×2, J1a, E-M123 (f1a, Ebla's node), R1b — **no P58** |
| Ebla (N. Syria) | 5 | 2600–1850 BCE | P58 ×3, E-Z830>M123, T1a |
| Oylum Höyük (Kilis) | 6 | 2500–1633 BCE | P58 ×2, T1a ×2, E-M123 ×2 (f1a~, Ebla's node) |
| **Alalakh / Tell Atchana** | **~20** | 1911–1428 BCE | **P58 ×6, J2a ×5, J2b ×2, T1a ×3, E-M78 (…a1a1c1, as at Akhziv), H2b, L1b, L2** |
| Hasanlu (NW Iran) | 5 | 2092–856 BCE | J1a ×2, J2a ×2, L1a |
| Dinkha Tepe (NW Iran) | 3 | 1821–1162 BCE | J2a ×2, P58 |

The chronology inside that table is itself a constraint. Arslantepe is J2a-dominant with **no P58 at all** through the fourth and third millennia; P58 enters the corridor at Ebla and Oylum Höyük around 2520–2500 BCE.

**The Caucasus is J1 country but not necessarily *P58 country.** Of the Caucasian J1 samples in the compilation, twenty-seven belong to J1a2b and only four are P58-derived, with J1b and unresolved J1 making up the remainder. The dense series — Taribana (Georgia, 3300 BCE), Balichi-Dzedzvebi (3000 BCE), Velikent (Dagestan, 2900 BCE), Dzhoghaz (Armenia, 2459–2375 BCE), Ginchi (Dagestan, 2320–2165 BCE) — is a diversification of the branch that is *not* P58. The clade's Caucasian depth and P58's southwestern turn are two different phenomena, and conflating them would repeat the sibling-branch



© 2026 Noël Alejandro Bota, J.D. — Educational, historical, and genealogical research use only (Fair Use). Drafting and editing assistance with Claude-AI. Verify all claims against primary sources.

error recorded above. While a few P58 samples *have been found/associated with ancient Armenia, most J1 linked to Hurro-Urartians belong to J1-Z1842. This cuts *both* ways: cannot be completely ruled out but cannot be *relied* upon.

**Three constraints any treatment of the corridor must accept.** Haplogroup **G is entirely absent** from this compilation, so claims about G2a in the Anatolian overlap require a different dataset. There is **no Hurrian reference population**: what exists is Bronze Age northern Levantine material from cities **that had Hurrian inhabitants and Hurrian personal names**, and affinity to those samples cannot be reported as affinity to "Hurrians" without committing the category error this document has spent its length prosecuting. And the archaeological hooks need staging — **Khabur Ware** is Middle Bronze Age northern Mesopotamian and its ethnic attribution is contested, while **Nuzi ware** belongs to the Mitanni period and carries the Hurrian association more securely.

The corridor is therefore left open rather than answered here. It is a positive thesis about a transmission zone; this document is a negative one about a category. The two do not share a spine and running them together would weaken both.



*© 2026 Noël Alejandro Bota, J.D. — Educational, historical, and genealogical research use only (Fair Use). Drafting and editing assistance with Claude-AI. Verify all claims against primary sources.*

## Sources

Agranat-Tamir, L., Waldman, S., Martin, M. A. S., Gokhman, D., Mishol, N., Eshel, T., Cheronet, O., Rohland, N., Mallick, S., Adamski, N., Lawson, A. M., Mah, M., Michel, M., Oppenheimer, J., Stewardson, K., Candilio, F., Keating, D., Gamarra, B., Tzur, S., . . . Reich, D. (2020). The genomic history of the Bronze Age Southern Levant. *Cell, 181*(5), 1146–1157.e11. https://doi.org/10.1016/j.cell.2020.04.024

Anthropic. (2026). *Claude* (Opus 5) [Large language model]. https://claude.ai

Chiaroni, J., King, R. J., Myres, N. M., Henn, B. M., Ducourneau, A., Mitchell, M. J., Boetsch, G., Sheikha, I., Lin, A. A., Nik-Ahd, M., Ahmad, J., Lattanzi, F., Herrera, R. J., Ibrahim, M. E., Brody, A., Semino, O., Kivisild, T., & Underhill, P. A. (2010). The emergence of Y-chromosome haplogroup J1e among Arabic-speaking populations. *European Journal of Human Genetics, 18*(3), 348–353. https://doi.org/10.1038/ejhg.2009.166

Cruciani, F., Trombetta, B., Sellitto, D., Massaia, A., Destro-Bisol, G., Watson, E., Beraud Colomb, E., Dugoujon, J.-M., Moral, P., & Scozzari, R. (2010). Human Y chromosome haplogroup R-V88: A paternal genetic record of early mid Holocene trans-Saharan connections and the spread of Chadic languages. *European Journal of Human Genetics, 18*(7), 800–807. https://doi.org/10.1038/ejhg.2009.231

Ehret, C., Keita, S. O. Y., & Newman, P. (2004). The origins of Afroasiatic. *Science, 306*(5702), 1680. https://doi.org/10.1126/science.306.5702.1680c

Grabbe, L. L. (Ed.). (2016). *The land of Canaan in the Late Bronze Age* (Library of Hebrew Bible / Old Testament Studies). Bloomsbury T&T Clark. [Includes Dijkstra, M., "Canaan in the transition from the Late Bronze to the Early Iron Age from an Egyptian perspective," and Grabbe, L. L., "Canaan under the rule of the Egyptian New Kingdom: From the Hyksos to the Sea Peoples."]

Haber, M., Doumet-Serhal, C., Scheib, C., Xue, Y., Danecek, P., Mezzavilla, M., Youhanna, S., Martiniano, R., Prado-Martinez, J., Szpak, M., Matisoo-Smith, E., Schutkowski, H., Mikulski, R., Zalloua, P., Kivisild, T., & Tyler-Smith, C. (2017). Continuity and admixture in the last five millennia of Levantine history from ancient Canaanite and present-day Lebanese genome sequences. *American Journal of Human Genetics, 101*(2), 274–282. https://doi.org/10.1016/j.ajhg.2017.06.013

Haber, M., Mezzavilla, M., Bergström, A., Prado-Martinez, J., Hallast, P., Saif-Ali, R., Al-Habori, M., Dedoussis, G., Zeggini, E., Blue-Smith, J., Wells, R. S., Xue, Y., Zalloua, P. A., & Tyler-Smith, C. (2016). Chad genetic diversity reveals an African history marked by multiple Holocene Eurasian migrations. *American Journal of Human Genetics, 99*(6), 1316–1324. https://doi.org/10.1016/j.ajhg.2016.10.012

*Haplogroup archaeogenetic record* [Data file]. (2026). Aggregated ancient Y-chromosome and mitochondrial haplogroup assignments for 3,180 published ancient individuals, with nested subclade labels, population labels, dates and site attributions. Supplied by the author. Note: subclade assignments in this compilation are re-derived rather than as-published, and are finer than the calls in the originating papers; individual assignments should be verified against the primary supplementary data before formal citation.

*Haplogroup E-M35*. (2026, October 25). In Wikipedia. https://en.wikipedia.org/wiki/Haplogroup_E-M35



*© 2026 Noël Alejandro Bota, J.D. — Educational, historical, and genealogical research use only (Fair Use). Drafting and editing assistance with Claude-AI. Verify all claims against primary sources.*

*Haplogroup H (Y-DNA)*. (2026, October 25). In Wikipedia. https://en.wikipedia.org/wiki/Haplogroup_H_(Y-DNA)

*Haplogroup J-M267*. (2026, October 25). In Wikipedia. https://en.wikipedia.org/wiki/Haplogroup_J-M267

*Haplogroup R1b*. (2026, October 25). In Wikipedia. https://en.wikipedia.org/wiki/Haplogroup_R1b

Harney, É., May, H., Shalem, D., Rohland, N., Mallick, S., Lazaridis, I., Sarig, R., Stewardson, K., Nordenfelt, S., Patterson, N., Hershkovitz, I., & Reich, D. (2018). Ancient DNA from Chalcolithic Israel reveals the role of population mixture in cultural transformation. *Nature Communications, 9*, Article 3336. https://doi.org/10.1038/s41467-018-05649-9

Harper, D. (n.d.). Semitism. In *Online Etymology Dictionary*. Retrieved October 25, 2026, from https://www.etymonline.com/word/semitism

Kitchen, A., Ehret, C., Assefa, S., & Mulligan, C. J. (2009). Bayesian phylogenetic analysis of Semitic languages identifies an Early Bronze Age origin of Semitic in the Near East. *Proceedings of the Royal Society B: Biological Sciences, 276*(1668), 2703–2710. https://doi.org/10.1098/rspb.2009.0408

Josephus, F. (1737). *The antiquities of the Jews* (W. Whiston, Trans.). Book 1, Chapter 6. (Original work composed ca. 94 CE). https://penelope.uchicago.edu/josephus/ant-1.html

Lazaridis, I., Nadel, D., Rollefson, G., Merrett, D. C., Rohland, N., Mallick, S., Fernandes, D., Novak, M., Gamarra, B., Sirak, K., Connell, S., Stewardson, K., Harney, É., Fu, Q., Gonzalez-Fortes, G., Jones, E. R., Roodenberg, S. A., Lengyel, G., Bocquentin, F., . . . Reich, D. (2016). Genomic insights into the origin of farming in the ancient Near East. *Nature, 536*(7617), 419–424. https://doi.org/10.1038/nature19310

Marcus, J. H., Posth, C., Ringbauer, H., Lai, L., Skeates, R., Sidore, C., Beckett, J., Furtwängler, A., Olivieri, A., Chiang, C. W. K., Al-Asadi, H., Dey, K., Joseph, T. A., Liu, C.-C., Der Sarkissian, C., Radzevičiūtė, R., Michel, M., Gradoli, M. G., Marongiu, P., . . . Novembre, J. (2020). Genetic history from the Middle Neolithic to present on the Mediterranean island of Sardinia. *Nature Communications, 11*, Article 939. https://doi.org/10.1038/s41467-020-14523-6

Militarev, A. (2009). Proto-Afrasian lexicon confirming West Asian homeland: Pastoralism. *Journal of Language Relationship, 1*, 95–106.

Sahakyan, H., Margaryan, A., Saag, L., Karmin, M., Flores, R., Haber, M., Kushniarevich, A., Khachatryan, Z., Bahmanimehr, A., Parik, J., Karafet, T., Yepiskoposyan, L., Reisberg, T., Metspalu, E., Behar, D. M., Yunusbayev, B., Balanovsky, O., Villems, R., & Rootsi, S. (2021). Origin and diffusion of human Y chromosome haplogroup J1-M267. *Scientific Reports, 11*, Article 6659. https://doi.org/10.1038/s41598-021-85883-2

Skourtanioti, E., Erdal, Y. S., Frangipane, M., Balossi Restelli, F., Yener, K. A., Pinnock, F., Matthiae, P., Özbal, R., Schoop, U.-D., Guliyev, F., Akhundov, T., Lyonnet, B., Hammer, E. L., Nugent, S. E., Burri, M., Neumann, G. U., Penske, S., Ingman, T., Akar, M., . . . Krause, J. (2020). Genomic history of Neolithic to Bronze Age Anatolia, Northern Levant, and Southern Caucasus. *Cell, 181*(5), 1158–1175.e28. https://doi.org/10.1016/j.cell.2020.04.044



*© 2026 Noël Alejandro Bota, J.D. — Educational, historical, and genealogical research use only (Fair Use). Drafting and editing assistance with Claude-AI. Verify all claims against primary sources.*

van Bekkum, K. (2018). The 'language of Canaan': Ancient Israel's history and the origins of Hebrew. In *Biblical Hebrew in context: Essays in Semitics and Old Testament texts in honour of Professor Jan P. Lettinga* (Oudtestamentische Studiën / Old Testament Studies, Vol. 74, pp. 67–88). Brill. https://doi.org/10.1163/9789004380851_008

## Cited through secondary sources — verify before formal use

The following were referenced in this document via summaries rather than the primary articles, and their full bibliographic details should be confirmed against the originals: Kulichová et al. (2017) on R1b-V88's non-specificity to Chadic speakers; Shriner and Rotimi on the Baggarization model; and the Chad genetic-history study summarized at https://scite.ai/reports/genetic-history-of-chad-K6YJxxM. The Tepe Hissar samples (I2512, I2514) are taken from the compilation; their originating publication is believed to be Narasimhan, V. M., et al. (2019), The formation of human populations in South and Central Asia, Science, 365(6457), eaat7487, https://doi.org/10.1126/science.aat7487 — confirm against its supplementary tables. The modern Iranian E-M78* reference comes from the Haplogroup E-M35 entry cited above.

Haplogroup assignments drawn from the archaeogenetic compilation are cited here by sample ID, site and date so that each can be traced to its originating publication. **A general warning about nested labels:** alphanumeric haplogroup labels encode tree position, not relatedness in time. Two labels differing in one character — J1a2a1a2 and J1a2a1b — can represent branches that diverged twelve thousand years ago. Any argument that treats adjacent labels as close relatives, or that reads a sibling branch's location as evidence for its sibling's, needs the divergence date stated explicitly before it can be assessed. The nested labels (for example E1b1b1b2a1a1a1a1f1a, J1a2a1a2d2b) are the compilation's own and do not appear in the source papers; the ISOGG equivalents given in the text (Z830, M123, M136, M78, V22, M293, P58, L620) were cross-checked against the ISOGG haplogroup trees, but the deep terminal assignments were not independently re-called and should be treated as provisional.

Author lists for the large-consortium genomics papers above follow the ellipsis convention for 21+ authors (first 19, ellipsis, final author). Confirm the intervening names against each publication before submission. The van Bekkum chapter appears in a Festschrift whose editors are not recorded here; APA requires them in the In-element, so add the editor names from the volume's title page before formal use.



*© 2026 Noël Alejandro Bota, J.D. — Educational, historical, and genealogical research use only (Fair Use). Drafting and editing assistance with Claude-AI. Verify all claims against primary sources.*