Algorithmic handwriting analysis of Judah's military correspondence sheds light on composition of biblical texts

Shira Faigenbaum-Golovin[a,1,2], Arie Shaus[a,1,2], Barak Sober[a,1,2], David Levin[a], Nadav Na'aman[b], Benjamin Sass[c], Eli Turkel[a], Eli Piasetzky[d], and Israel Finkelstein[c]

[a]Department of Applied Mathematics, Sackler Faculty of Exact Sciences, Tel Aviv University, Tel Aviv 69978, Israel; [b]Department of Jewish History, Tel Aviv University, Tel Aviv 69978, Israel; [c]Jacob M. Alkow Department of Archaeology and Ancient Near Eastern Civilizations, Tel Aviv University, Tel Aviv 69978, Israel; and [d]School of Physics and Astronomy, Sackler Faculty of Exact Sciences, Tel Aviv University, Tel Aviv 69978, Israel

Edited by Klara Kedem, Ben-Gurion University, Be'er Sheva, Israel, and accepted by the Editorial Board March 3, 2016 (received for review November 17, 2015)

The relationship between the expansion of literacy in Judah and composition of biblical texts has attracted scholarly attention for over a century. Information on this issue can be deduced from Hebrew inscriptions from the final phase of the first Temple period. We report our investigation of 16 inscriptions from the Judahite desert fortress of Arad, dated *ca*. 600 BCE—the eve of Nebuchadnezzar's destruction of Jerusalem. The inquiry is based on new methods for image processing and document analysis, as well as machine learning algorithms. These techniques enable identification of the minimal number of authors in a given group of inscriptions. Our algorithmic analysis, complemented by the textual information, reveals a minimum of six authors within the examined inscriptions. The results indicate that in this remote fort literacy had spread throughout the military hierarchy, down to the quartermaster and probably even below that rank. This implies that an educational infrastructure that could support the composition of literary texts in Judah already existed before the destruction of the first Temple. A similar level of literacy in this area is attested again only 400 y later, *ca*. 200 BCE.

biblical exegesis | literacy level | Arad ostraca | document analysis | machine learning

Based on biblical exegesis and historical considerations scholars debate whether the first major phase of compilation of biblical texts in Jerusalem took place before or after the destruction of the city by the Babylonians in 586 BCE (e.g., ref. 1). A related—and also disputed—issue is the level of literacy, that is, the basic ability to communicate in writing, especially in the Hebrew kingdoms of Israel and Judah (2). The best way to answer this question is to look at the material evidence: the corpus of inscriptions that originated from archaeological excavations (e.g., ref. 3). Inscriptions citing biblical texts, or related to them, are rarely found (for two Jerusalem amulets possibly dating to this period, echoing the priestly blessing in Numbers 6:23–26, see refs. 4 and 5), probably because papyrus and parchment are not well preserved in the climate of the region. However, ostraca (inscriptions in ink on ceramic sherds) that deal with more mundane issues can also shed light on the volume and quality of writing and on the recognition of the power of the written word in the society.

To explore the degree of literacy and stage setting for compilation of literary texts in monarchic Judah, we turned to Hebrew ostraca from the final days of the kingdom, before its destruction by Nebuchadnezzar in 586 BCE and the deportation of its elite to Babylonia. Several corpora of inscriptions exist for this period. We focused on the corpus of over 100 Hebrew ostraca found at the fortress of Arad, located in arid southern Judah, on the border of the kingdom with Edom (see ref. 6 and Fig. 1). The inscriptions contain military commands regarding movement of troops and provision of supplies (wine, oil, and flour) set against the background of the stormy events of the final years before the fall of Judah. They include orders that came to the fortress of Arad from higher echelons in the Judahite military system, as well as correspondence with neighboring forts. One of the inscriptions mentions "the King of Judah" and another "the house of YHWH," referring to the Temple in Jerusalem. Most of the provision orders that mention the *Kittiyim*—apparently a Greek mercenary unit (7)—were found on the floor of a single room. They are addressed to a person named Eliashib, the quartermaster in the fortress. It has been suggested that most of Eliashib's letters involve the registration of about one month's expenses (8).

Of all of the corpora of Hebrew inscriptions, Arad provides the best set of data for exploring the question of literacy at the end of the first Temple period: (*i*) The lion's share of the corpus represents a short time span of a few years *ca*. 600 BCE; (*ii*) it comes from a remote region of the kingdom, where the spread of literacy is more significant than its dissemination in the capital; and (*iii*) it is connected to Judah's military administration and hence bureaucratic apparatus. Identifying the number of "hands" (i.e., authors) involved in this corpus can shed light on the

Significance

Scholars debate whether the first major phase of compilation of biblical texts took place before or after the destruction of Jerusalem in 586 BCE. Proliferation of literacy is considered a precondition for the creation of such texts. Ancient inscriptions provide important evidence of the proliferation of literacy. This paper focuses on 16 ink inscriptions found in the desert fortress of Arad, written *ca*. 600 BCE. By using novel image processing and machine learning algorithms we deduce the presence of at least six authors in this corpus. This indicates a high degree of literacy in the Judahite administrative apparatus and provides a possible stage setting for compilation of biblical texts. After the kingdom's demise, a similar literacy level reemerges only *ca*. 200 BCE.

Author contributions: S.F.-G., A.S., and B. Sober designed research; S.F.-G., A.S., and B. Sober performed research; S.F.-G., A.S., and B. Sober contributed new reagents/analytic tools; D.L. and E.T. supervised the development of the algorithms; N.N., B. Sass, and I.F. provided archaeological and epigraphical analysis and historical reconstruction; E.P. supervised the development of the algorithms; S.F.-G., A.S., and B. Sober analyzed data; S.F.-G., A.S., B. Sober, D.L., N.N., B. Sass, E.T., E.P., and I.F. wrote the paper; and E.P. and I.F. headed the research team.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission. K.K. is a guest editor invited by the Editorial Board.

Data deposition: Two datasets are provided on our institutional website, with free and open access: www-nuclear.tau.ac.il/~eip/ostraca/DataSets/Modern_Hebrew.zip and www-nuclear.tau.ac.il/~eip/ostraca/DataSets/Arad_Ancient_Hebrew.zip.

[1]S.F.-G., A.S., and B. Sober contributed equally to this work.

[2]To whom correspondence may be addressed. Email: shirafaigen@gmail.com, ashaus@post.tau.ac.il, or baraksov@post.tau.ac.il.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1522200113/-/DCSupplemental.

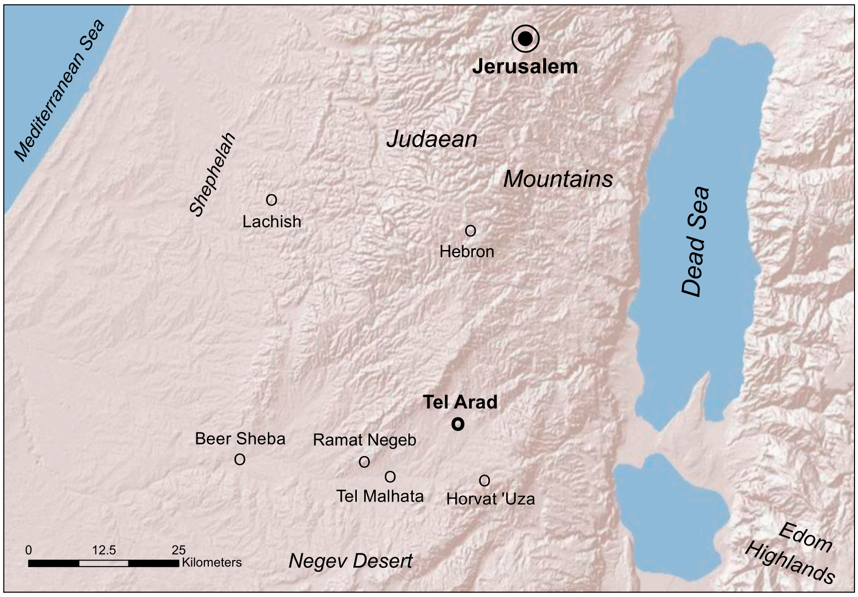

Fig. 1. Main towns in Judah and sites in the Beer Sheba Valley mentioned in the article.

dissemination of writing, and consequently on the spread of literacy in Judah.

Algorithmic Apparatus

One might try to use existing computerized algorithms for automatic handwriting comparison purposes. However, an algorithmic analysis of the Arad corpus via readily available means is hampered by several factors. First, the poor state of preservation of the ostraca (Fig. 2) could not be remedied by existing image acquisition methods (9, 10). Second, the imperfect digital images present a challenge for image segmentation and enhancement methods (11, 12). Finally, recognizing hands via document analysis algorithms is a tantalizing problem even in a modern writing setting (13). Consequently, we developed new methods for image processing and document analysis, as well as machine learning algorithms. These techniques allow us to identify the minimal number of authors represented in a given group of ostraca.

Our algorithmic sequence consisted of three consecutive stages, operating on digital images of the ostraca (see *Supporting Information*). All of the stages are fully automatic, with the exception of the first, which is a semiautomatic step.

i) Restoring characters (see example in Fig. 3; also see *Supporting Information* and ref. 14)

ii) Extraction of characters' features, describing their different aspects (e.g., angles between strokes and character profiles), and measuring the similarity ("distances") between the characters' feature vectors.

iii) Testing the null hypothesis H_0 (for each pair of ostraca), that two given inscriptions were written by the same author. A corresponding P value (P) is deduced, leveraging the data from the previous step. If $P \leq 0.2$, we reject H_0 and accept the competing hypothesis of two different authors; otherwise, we remain undecided.

The end product is a table containing the P for a comparison of each pair of ostraca. Before implementing our methodology on the Arad corpus, it was thoroughly tested on modern Hebrew handwritings and found solid (see *Supporting Information* for details).

Results

Using this computerized procedure we analyzed 16 inscriptions from the Arad fortress (namely, ostraca 1, 2, 3, 5, 7, 8, 16, 17, 18,

Fig. 2. Ostraca from Arad (see ref. 6): numbers 24 (*A*), 5 (*B*), and 40 (*C*). The poor state of preservation, including stains, erased characters, and blurred text, is evident. Images are courtesy of the Institute of Archaeology, Tel Aviv University, and of the Israel Antiquities Authority.

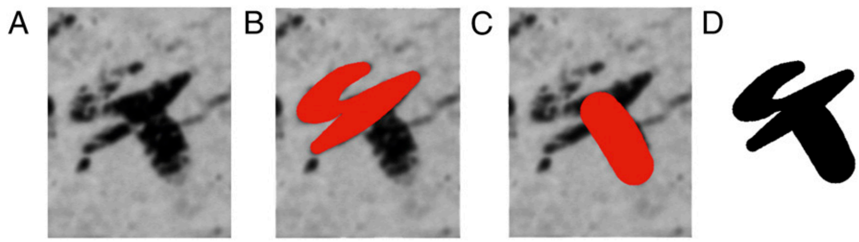

Fig. 3. Restoration of the character *waw* in Arad ostracon 24 (see ref. 14). (*A*) The original image. (*B* and *C*) reconstructed strokes. (*D*) The resulting character restoration (see *Supporting Information* for further details). Images are courtesy of the Institute of Archaeology, Tel Aviv University, and of the Israel Antiquities Authority.

21, 24, 31, 38, 39, 40, and 111), which are relatively legible and have a sufficient number of characters for examination. Two of the inscriptions (ostraca 17 and 39) are inscribed on both sides of the sherd, bringing the number of texts under investigation to 18. The results are summarized in Table 1. The ostraca numbers head the rows and columns of the table, with the intersection cells providing the comparisons' P. The cells with $P \leq 0.2$ are marked in red, indicating that the two ostraca are considered to be written by different authors. We reiterate that when $P > 0.2$ we cannot claim that they were written by a single author.

The results allow us to estimate the minimal number of writers in the tested inscriptions. For example, the examination of ostraca 7, 18, 24, and 40 reveals that their authors are pairwise distinct; in fact, six such "quadruplets" can be identified in Table 1, rendering the existence of at least four authors as highly likely; see *Supporting Information* for details. Therefore, based on the statistical analysis, it can be deduced that there are at least four unique hands in the tested corpus. Our algorithmic observations can be further supplemented by the textual and archaeological context of the ostraca, deliberately avoided until this point. In particular, the prosaic lists of names in ostraca 31 and 39[*] were most likely composed at Arad, as opposed to ostraca 7, 18, 24, and 40, which were probably dispatched from other locations.[†] As per the table, ostracon 31 differs from both sides of ostracon 39; we can thus conjecture an existence of two additional authors, totaling at least six distinct writers.

Discussion

Identifying the military ranks of the authors can provide information regarding the spread of literacy within the Judahite army. Our proposed reconstruction of the hierarchical relations between the signees and the addressees of the examined inscriptions is as follows[‡] (see Fig. 4):

i) The King of Judah: mentioned in ostracon 24 as dictating the overall military strategy
ii) An unnamed military commander: the author of ostracon 24
iii) Malkiyahu, the commander of the Arad fortress: mentioned in ostracon 24 and the recipient of ostracon 40[§]
iv) Eliashib, the quartermaster of the Arad fortress: the addressee of ostraca 1–16 and 18; mentioned in ostracon 17a; the writer of ostracon 31
v) Eliashib's subordinate: addressing Eliashib as "my lord" in ostracon 18

Following this reconstruction, it is reasonable to deduce the proliferation of literacy among the Judahite army ranks *ca*. 600 BCE. A contending claim that the ostraca were written by professional scribes can be dismissed with two arguments: the existence of two distinct writers in the tiny fortress of Arad (authors of ostraca 31 and 39) and the textual content of the inscriptions: Ostracon 1 orders the recipient (Eliashib) "write the name of the day," ostracon 7 commands "and write it before you…," and in ostracon 40 (reconstructions in refs. 6 and 18) the author mentions that he had written the letter. Thus, rather than implying the existence of scribes accompanying every Judahite official, the written evidence suggests a high degree of literacy in the entire Judahite chain of command.

The dissemination of writing within the Judahite army around 600 BCE is also confirmed by the existence of other military-related corpora of ostraca, at Horvat ʿUza (19) and Tel Malḥata (20) in the vicinity of Arad, and at Lachish[¶] in the Shephelah (summary in ref. 3)—all located on the borders of Judah (Fig. 1). We assume that in all these locations the situation was similar to Arad, with even the most mundane orders written down occasionally. In other words, the entire army apparatus, from high-ranking officials to humble vice-quartermasters of small desert outposts far from the center, was literate, in the sense of the ability to communicate in writing.

To support this bureaucratic apparatus, an appropriate educational system must have existed in Judah at the end of the first Temple period (2, 21–23). Additional evidence supporting writing awareness by the lowest echelons of society seems to come from the Meẓad Hashavyahu ostracon (24), which contains a complaint by a corvée worker against one of his overseers (most scholars agree that it was composed with the aid of a scribe).

Extrapolating the minimum of six authors in 16 Arad ostraca to the entire Arad corpus, to the whole military system in the southern Judahite frontier, to military posts in other sectors of the kingdom, to central administration towns such as Lachish, and to

[*] Contrary to the excavator's association of ostraca 31 and 39 with Stratum VII (ref. 6, also ref. 15) rather than VI where most of the examined ostraca were found, we agree with critics (16, 17) that these strata are in fact one and the same. Note that ostracon 31 was found in locus 779, alongside three seals of Eliashib (the addressee of ostraca 1–16 and 18, from Strata VI).

[†] Ostraca 5, 7, 17a, 18, and 24 were most probably written in other locations (6). Ostracon 40 may have been written by troop commanders Gemaryahu and Nehemyahu (see the following note) with some ties to Arad fortress; their names also appear at ostracon 31. This renders the common authorship of ostraca 31 and 40 unlikely. Furthermore, from Table 1, ostraca 40 and 39a have different authors.

[‡] We conjecture that the status of the officers who commanded the supplies to the Kittiyim (the Greek or Cypriot mercenary unit), who wrote ostraca 1–8 and 17a, was similar to that of Malkiyahu (the commander of the fortress at Arad), and in any case they were Eliashib's superiors. Also note that Gemaryahu and Nehemyahu (ostracon 40) are Malkiyahu's subordinates, whereas Hananyahu (author of ostracon 16, also mentioned in ostracon 3) is probably Eliashib's counterpart in Beer Sheba. The textual content of the ostraca also suggests differentiation between combatant and logistics-oriented officials (Fig. 4).

[§] Contrary to the excavator's dating of ostracon 40 to Stratum VIII of the late 8th century (ref. 6, also ref. 17), it should probably be placed a century later, along with ostracon 24 (see ref. 18 for details). Note that a conflict between the vassal kingdoms of Judah and Edom, seemingly hinted at in this inscription, is unlikely under the strong rule of the Assyrian empire in the region (*ca*. 730–630 BCE), especially along the vitally important Arabian trade routes.

[¶] In fact, Lachish ostracon 3, also containing military correspondence, represents the most unambiguous evidence of a writing officer. The author seems offended by a suggestion that he is assisted by a scribe. See detail, including discussion regarding the literacy of army personnel, in ref. 2.

Table 1. Comparison between different Arad ostraca

No.	Ostraca content	1	2	3	5	7	8	16	17a	17b	18	21	24	31	38	39a	39b	40	111
1	Order to Eliashib, supply of provisions for the Kittiyim		0.64	0.50	0.91	0.30	0.64	0.51	0.98	0.78	0.53	0.24	0.003	0.10	0.27	0.41	0.06	0.23	0.79
2	Order to Eliashib, supply of provisions for the Kittiyim	0.64		1.00	1.00	0.72	1.00	0.39	0.85	0.78	0.31	0.75	0.79	0.06	0.38	0.98	0.70	0.11	0.96
3	Order to Eliashib mentioning Hananyahu, concerning provisions to Beer Sheba	0.50	1.00		0.23	0.06	0.55	0.36	1.00	0.77	0.27	0.94	0.72	0.16	0.61	0.96	0.84	0.22	0.79
5	Order to Eliashib, supply of provisions, probably for the Kittiyim	0.91	1.00	0.23		0.53	0.60	0.60	0.19	0.40	0.07	0.46	0.12	0.01	0.40	0.24	0.21	0.07	0.98
7	Order to Eliashib, supply of provisions for the Kittiyim	0.30	0.72	0.06	0.53		0.03	0.76	0.17	0.48	0.004	0.43	0.05	0.07	0.27	0.35	1.00	0.15	0.05
8	Order to Eliashib, supply of provisions for the Kittiyim	0.64	1.00	0.55	0.60	0.03		0.68	0.07	1.00	0.17	0.33	0.74	0.42	0.20	0.67	1.00	1.00	0.93
16	Letter to Eliashib from Hananyahu	0.51	0.39	0.36	0.60	0.76	0.68		0.33	1.00	0.03	0.80	0.13	0.38	0.38	0.41	0.40	0.72	0.68
17a	Order to Nahum to proceed to the house of Eliashib to collect provisions	0.98	0.85	1.00	0.19	0.17	0.07	0.33		1.00	0.92	0.36	0.13	0.41	1.00	0.68	1.00	0.17	0.68
17b	Note that Nahum provided provisions to the Kittiyim	0.78	0.78	0.77	0.40	0.48	1.00	1.00	1.00		1.00	0.35	0.40	0.47	1.00	1.00	0.33	0.20	0.40
18	Report to Eliashib from a subordinate fulfilling an order; mention of the Temple	0.53	0.31	0.27	0.07	0.004	0.17	0.03	0.92	1.00		3×10^{-4}	0.02	0.20	0.32	0.94	0.86	0.04	0.73
21	Letter to Gedalyahu from a subordinate, Yehokal	0.24	0.75	0.94	0.46	0.43	0.33	0.80	0.36	0.35	3×10^{-4}		0.35	0.04	0.23	0.71	0.21	0.31	0.90
24	A royal decree ordering the reinforcement of Ramat Negeb against Edom	0.003	0.79	0.72	0.12	0.05	0.74	0.13	0.13	0.40	0.02	0.35		0.01	0.05	0.73	0.38	0.002	0.92
31	List of names	0.10	0.06	0.16	0.01	0.07	0.42	0.38	0.41	0.47	0.20	0.04	0.01		0.33	0.16	0.11	0.35	0.57
38	List of names (including the son of Eliashib)	0.27	0.38	0.61	0.40	0.27	0.20	0.38	1.00	1.00	0.32	0.23	0.05	0.33		0.77	0.33	0.70	0.77
39a	List of names	0.41	0.98	0.96	0.24	0.35	0.67	0.41	0.68	1.00	0.94	0.71	0.73	0.16	0.77		1.00	0.04	0.75
39b	List of names	0.06	0.70	0.84	0.21	1.00	1.00	0.40	1.00	0.33	0.86	0.21	0.38	0.11	0.33	1.00		0.42	0.42
40	Gemaryahu and Nehemyahu report to Malkiyahu mentioning Edom and the king of Judah	0.23	0.11	0.22	0.07	0.15	1.00	0.72	0.17	0.20	0.04	0.31	0.002	0.35	0.70	0.04	0.42		0.67
111	Fragmentary, mentioning guard and horses	0.79	0.96	0.79	0.98	0.05	0.93	0.68	0.68	0.40	0.73	0.90	0.92	0.57	0.77	0.75	0.42	0.67	

A $P \leq 0.2$, highlighted in red, indicates rejection of "single writer" hypothesis, hence accepting a "two different authors" alternative. Note that ostraca 17 and 39 contain writing on both sides of the sherd (marked as "a" and "b").

the capital, Jerusalem, a significant number of literate individuals can be assumed to have lived in Judah *ca*. 600 BCE.

The spread of literacy in late-monarchic Judah provides a possible stage setting for the compilation of literary works. True, biblical texts could have been written by a few and kept in seclusion in the Jerusalem Temple, and the illiterate populace could have been informed about them in public readings and verbal messages by these few (e.g., 2 Kings 23:2, referring to the period discussed here). However, widespread literacy offers a better background for the composition of ambitious works such as the Book of Deuteronomy and the history of Ancient Israel in the Books of Joshua to Kings (known as the Deuteronomistic History), which formed the platform for Judahite ideology and theology (e.g., ref. 25). Ideally, to deduce from literacy on the composition of literary (to differ from mundane) texts, we should have conducted comparative research on the centuries after the destruction of Jerusalem, a period when other biblical texts were written in both Jerusalem and Babylonia according to current textual research (e.g., refs. 1 and 26). However, in the Babylonian, Persian, and early Hellenistic periods, Jerusalem and the southern highlands show almost no evidence in the form of Hebrew inscriptions. In fact, not a single securely dated Hebrew inscription has been found in this territory for the period between 586 and *ca*. 350 BCE[#]—not an ostracon or a seal, a seal impression, or a bulla [the little that we know of this period is in Aramaic, the script of the newly present Persian empire (27)]. This should come as no surprise, because the destruction of Judah brought about the collapse of the kingdom's bureaucracy and deportation of many of the literati. Still, for the centuries between *ca*. 600 and 200 BCE, the tension between current biblical exegesis (arguing for massive composition of texts) and the negative archaeological evidence remains unresolved.

Materials and Methods

This research was conducted on two datasets of written material. The main document assemblage was a corpus of 16 Hebrew ostraca inscriptions found at the Arad fortress (*ca*. 600 BCE). The research was performed on digital

[#] A few coins with Hebrew characters do appear between *ca*. 350 and 200 BCE.

ANTHROPOLOGY

APPLIED MATHEMATICS

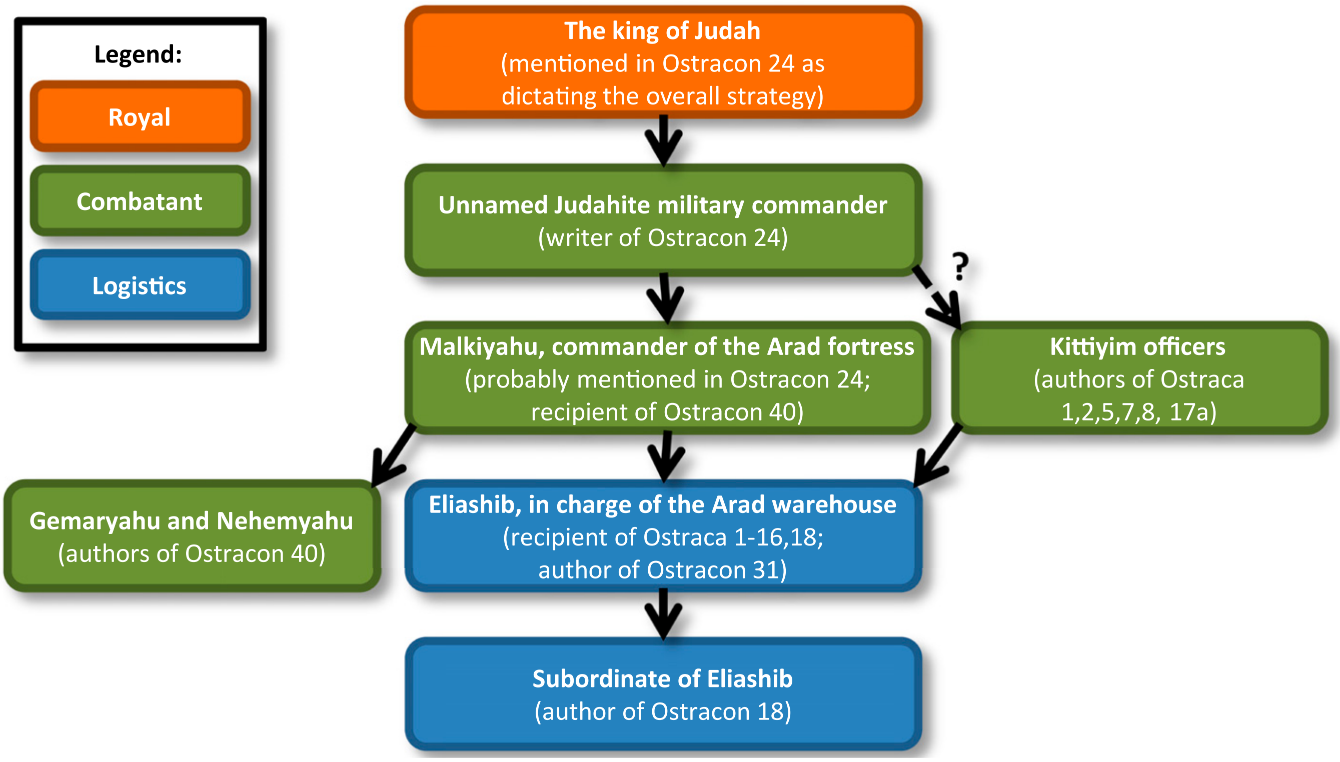

Fig. 4. Reconstruction of the hierarchical relations between authors and recipients in the examined Arad inscriptions; also indicated is the differentiation between combatant and logistics officials.

images of these inscriptions. A second dataset, used to validate the algorithm, contained handwriting samples collected from 18 present-day writers of Modern Hebrew.

The aim of our main algorithm was to differentiate between writers in a given set of texts. This algorithm consisted of several stages. In the first step, character restoration, the image of the inscription was segmented into (often noisy) characters that were restored via a semiautomatic reconstruction procedure. The method was based on the representation of a character as a union of individual strokes that were treated independently and later recombined. The purpose of stroke restoration was to imitate a reed pen's movement using several manually sampled key points. An optimization of the pen's trajectory was performed for all intermediate sampled points. The restoration was conducted via the minimization of image energy functional, which took into account the adherence to the original image, the smoothness of the stroke, as well as certain properties of the reed radius. The minimization problem was solved by performing gradient descent iterations on a cubic-spline representation of the stroke. The end product of the reconstruction was a binary image of the character, incorporating all its strokes (see Figs. S1 and S2).

The second stage of the algorithm, letter comparison, relied on features extracted from the characters' binary images, used to automatically compare characters from different texts. Several features were adapted, referring to aspects such as the character's overall shape, the angles between strokes, the character's center of gravity, as well as its horizontal and vertical projections. The features in use were SIFT (28), Zernike (29), DCT, K_d-tree (30), Image projections (31), L_1, and CMI (32). Additionally, for each feature, a respective distance was defined. Later on, all these distances were combined into a single, generalized feature vector. This vector described each character by the degree of its proximity to all of the characters, using all of the features. Finally, a distance between any two characters was calculated according to the Euclidean distance between their generalized feature vectors (see Table S1 for details concerning various features in use).

The final stage of the algorithm addressed the main question, What is the probability that two given texts were written by the same author? This was achieved by posing an alternative null hypothesis H_0 ("both texts were written by the same author") and attempting to reject it by conducting a relevant experiment. If its outcome was unlikely ($P \leq 0.2$), we rejected the H_0 and concluded that the documents were written by two individuals. Alternatively, if the occurrence of H_0 was probable ($P > 0.2$), we remained agnostic.

The experiment testing the H_0 performed a clustering on a set of letters from the two tested inscriptions (of specific type, e.g., *alep*[||]), disregarding their affiliation to either of the inscriptions. The clustering results should have resembled the original inscriptions if two different writers were present, while being random if this was not the case. Although this kind of test could have been performed on one specific letter, we could gain additional statistical significance if several different letters (e.g., *alep*, *he*, *waw*, etc.) were present in the compared documents. Subsequently, several independent experiments were conducted (one for each letter), and their P values were combined via the well-established Fisher's method (33). The combination represented the probability that H_0 was true based on all of the evidence at our disposal (see Fig. S3 for an illustration of the procedure's flow).

See *Supporting Information* for additional details regarding the methods in use and their results on both Ancient and Modern Hebrew datasets (available at www-nuclear.tau.ac.il/~eip/ostraca/DataSets/Arad_Ancient_Hebrew.zip and www-nuclear.tau.ac.il/~eip/ostraca/DataSets/Modern_Hebrew.zip, respectively). In particular, see Figs. S4 and S5 for samples taken from Modern and Ancient Hebrew datasets, respectively. Additionally, Table S2 summarizes the results of the Modern Hebrew experiment, while Table S3 provides statistics regarding the characters utilized in the Ancient Hebrew experiment.

ACKNOWLEDGMENTS. This research was made possible by the dedicated work of Ms. Ma'ayan Mor. The kind assistance of Dr. Shirly Ben-Dor Evian, Ms. Sivan Einhorn, Ms. Noa Evron, Dr. Anat Mendel, Ms. Myrna Pollak, Mr. Michael Cordonsky, and Mr. Assaf Kleiman is greatly appreciated. We also thank the PNAS editor and the reviewers for their helpful comments and suggestions. A.S. thanks the Azrieli Foundation for the award of an Azrieli Fellowship. Ostracon images are courtesy of the Institute of Archaeology, Tel Aviv University, and of the Israel Antiquities Authority. The research reported here received initial funding from the Israel Science Foundation – F.I.R.S.T. (Bikura) Individual Grant 644/08, as well as Israel Science Foundation Grant

[||]The Latin transliteration of the letter names differs slightly between Modern and Ancient Hebrew. For Ancient Hebrew, several spellings can be found in the literature: *alep/aleph, bet, gimel, dalet, he, waw, zayin, het/ḥet, tet/ṭet, yod, kap/kaf, lamed, mem, nun, samek/samekh, ayin/ʿayin, pe, sade/ṣade, qop/qof, resh, shin, taw*. For Modern Hebrew, the Unicode standard names are *alef, bet, gimel, dalet, he, vav, zayin, het, tet, yod, kaf, lamed, mem, nun, samekh, ayin, pe, tsadi, qof, resh, shin, tav*. For simplicity's sake, in what follows, we use the first orthography (without the diacritics) for each letter.

1457/13. The research was also funded by the European Research Council under the European Community's Seventh Framework Programme (FP7/2007-2013)/ERC Grant Agreement 229418, and by an Early Israel grant (New Horizons project), Tel Aviv University. This study was also supported by a generous donation from Mr. Jacques Chahine, made through the French Friends of Tel Aviv University.

1. Schmid K (2012) *The Old Testament: A Literary History* (Fortress, Minneapolis).
2. Rollston CA (2010) *Writing and Literacy in the World of Ancient Israel: Epigraphic Evidence from the Iron Age* (Society of Biblical Literature, Atlanta).
3. Aḥituv S (2008) *Echoes from the Past: Hebrew and Cognate Inscriptions from the Biblical Period* (Carta, Jerusalem).
4. Barkay G (1992) The priestly benediction on silver plaques from Ketef Hinnom in Jerusalem. *Tel Aviv* 19(2):139–192.
5. Barkay G, Vaughn AG, Lundberg MJ, Zuckerman B (2004) The amulets from Ketef Hinnom: A new edition and evaluation. *Bull Am Schools Orient Res* 334: 41–71.
6. Aharoni Y (1981) *Arad Inscriptions* (Israel Exploration Society, Jerusalem).
7. Na'aman N (2011) Textual and historical notes on the Eliashib archive from Arad. *Tel Aviv* 38(1):83–93.
8. Lemaire A (1977) *Inscriptions Hébraïques, Vol. 1: Les Ostraca*. Littératures anciennes du Proche-Orient 9 (Edicions du Cerf, Paris), pp 230–231.
9. Faigenbaum-Golovin S, et al. (2015) Computerized paleographic investigation of Hebrew Iron Age ostraca. *Radiocarbon* 57(2):317–325.
10. Faigenbaum S, et al. (2012) Multispectral images of ostraca: Acquisition and analysis. *J Archaeol Sci* 39(12):3581–3590.
11. Shaus A, Turkel E, Piasetzky E (2012) Binarization of First Temple Period inscriptions - performance of existing algorithms and a new registration based scheme. Proceedings of the 13th International Conference on Frontiers in Handwriting Recognition (IEEE Computer Society, Los Alamitos, CA), pp 641–646.
12. Shaus A, Sober B, Turkel E, Piasetzky E (2013) Improving binarization via sparse methods. Proceedings of the 16th International Graphonomics Society Conference (Tokyo University of Agriculture and Technology Press, Tokyo), pp 163–166.
13. Louloudis G, Gatos B, Stamatopoulos N (2012) ICFHR 2012 competition on writer identification challenge 1: Latin/Greek documents. Proceedings of the 13th International Conference on Frontiers in Handwriting Recognition (IEEE Computer Society, Los Alamitos, CA), pp 829–834.
14. Sober B, Levin D (2016) Computer aided restoration of handwritten character strokes. arXiv:1602.07038.
15. Herzog Z (2002) The fortress mound at Tel Arad: An interim report. *Tel Aviv* 29(1): 3–109.
16. Mazar A, Netzer E (1986) On the Israelite fortress at Arad. *Bull Am Schools Orient Res* 263:87–91.
17. Ussishkin D (1988) The date of the Judean shrine at Arad, Israel. *Explor J* 38:142–157.
18. Na'aman N (2003) Ostracon 40 from Arad reconsidered. *Saxa loquentur. Studien zur Archäologie Palästinas/Israels. Festschrift für Volkmar Fritz zum 65 Geburtstag*, eds den Hertog CG, Hübner U, Münger S (Ugarit, Münster, Germany), pp 199–204.
19. Beit-Arieh I (2007) *Horvat 'Uza and Horvat Radum: Two Fortresses in the Biblical Negev*. Tel Aviv University Monograph Series 25 (Tel Aviv Univ, Tel Aviv).
20. Beit-Arieh I, Freud L (2015) *Tel Malḥata: A central city in the biblical Negev*. Tel Aviv University Monograph Series 32 (Tel Aviv Univ, Tel Aviv).
21. Rollston CA (1999) The script of Hebrew Ostraca of the Iron Age: 8th–6th centuries BCE. PhD thesis (Johns Hopkins Univ, Baltimore).
22. Rollston CA (2006) Scribal education in ancient Israel: The Old Hebrew epigraphic evidence. *Bull Am Schools Orient Res* 344:47–74.
23. Lemaire A (1981) *Les écoles et la formation de la Bible dans l'ancien Israël*. Orbis Biblicus et Orientalis 39 (Editions Universitaires, Fribourg, Switzerland).
24. Naveh J (1960) A Hebrew letter from the seventh century B.C. *Isr Explor J* 10(3): 129–139.
25. Na'aman N (2002) *The Past that Shapes the Present: The Creation of Biblical Historiography in the Late First Temple Period and After the Downfall* (Bialik Institute, Jerusalem). Hebrew.
26. Albertz R (2003) *Israel in Exile: The History and Literature of the Sixth Century B.C.E* (Society of Biblical Literature, Atlanta).
27. Lipschits O, Vanderhooft DS (2011) *The Yehud Stamp Impressions: A Corpus of Inscribed Impressions from the Persian and Hellenistic Periods in Judah* (Eisenbrauns, Winona Lake, IN).
28. Lowe DG (2004) Distinctive image features from scale-invariant keypoints. *Int J Comput Vis* 60(2):91–110.
29. Tahmasbi A, Saki F, Shokouhi SB (2011) Classification of benign and malignant masses based on Zernike moments. *Comput Biol Med* 41(8):726–735.
30. Sexton A, Todman A, Woodward K (2000) Font recognition using shape-based quad-tree and kd-tree decomposition, Proceedings of the 3rd International Conference on Computer Vision, Pattern Recognition and Image Processing (IEEE Computer Society, Los Alamitos, CA), pp 212–215.
31. Trier ØD, Jain AK, Taxt T (1996) Feature extraction methods for character recognition—A survey. *Pattern Recognit* 29(4):641–662.
32. Shaus A, Turkel E, Piasetzky E (2012) Quality evaluation of facsimiles of Hebrew First Temple period inscriptions. Proceedings of the 10th IAPR International Workshop on Document Analysis Systems (IEEE Computer Society, Los Alamitos, CA), pp 170–174.
33. Fisher RA (1925) *Statistical Methods for Research Workers* (Oliver and Boyd, Edinburgh).
34. Panagopoulos M, Papaodysseus C, Rousopoulos P, Dafi D, Tracy S (2009) Automatic writer identification of ancient Greek inscriptions. *IEEE Trans Pattern Anal Mach Intell* 31(8):1404–1414.
35. Mumford D, Shah J (1989) Optimal approximations by piecewise smooth functions and associated variational problems. *Commun Pure Appl Math* 42(5):577–685.
36. Kass M, Witkin A, Terzopoulos D (1988) Snakes: Active contour models. *Int J Comput Vis* 1(4):321–331.
37. Freund Y, Schapire RE (1997) A decision-theoretic generalization of on-line learning and an application to boosting. *J Comput Syst Sci* 55(1):119–139.
38. Sivic J, Zisserman A (2003) Video Google: A text retrieval approach to object matching in videos. Proceedings of the 9th International Conference on Computer Vision (IEEE Computer Society, Los Alamitos, CA), pp 1470–1477.
39. Tahmasbi A (2012) Zernike moments. Available at www.mathworks.com/matlabcentral/fileexchange/38900-zernike-moments.
40. Armon S (2012) Descriptor for shapes and letters (feature extraction). Available at www.mathworks.com/matlabcentral/fileexchange/35038-descriptor-for-shapes-and-letters-feature-extraction.