

Toward a Decorative Grammar of Chinese Lacquerware in the National Palace Museum Taipei

Peng Yue

yp6253344@163.com

Kangwon National University

Xiao Zhang

Christian-Albrechts-Universität zu Kiel

Jia Hou

Zhengzhou Business University

Article

Keywords: Chinese lacquerware, decorative grammar, motif analysis, spatial relations, computational heritage, museum collections, inter-coder reliability

Posted Date: October 20th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-11195742/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

Abstract

Decorative motifs are often catalogued as isolated elements, although their organization also depends on recurring combinations and spatial relations. We analysed 509 Chinese lacquerware objects from the National Palace Museum, Taipei using an evidence-aware, relation-based workflow. The adjudicated master contains 71 motif classes and 2,054 canonical relation triplets. Recurrent configurations included KEY_FRET[BORDER]-BORDER_OF->OBJECT_COMPOSITION (26.9%) and CLOUD[STRUCTURAL]-SURROUNDS->DRAGON[PRIMARY] (9.4%). Excluding 56 adjudicated objects retained 19 of the 20 highest-weight triplets (Spearman rank correlation = 0.790). Ming-Qing comparisons identified differences in selected motif prevalences after Holm and Benjamini-Hochberg adjustment, although chronological coverage was uneven. Because the pre-adjudication reliability gate remained HOLD, the dataset is interpreted descriptively rather than as independently reliable ground truth. The workflow operationalizes decorative grammar while keeping historical and iconographic interpretation separate from observable structure.

Introduction

Chinese lacquerware combines vessel form, surface technique and pictorial or ornamental programs in ways that are difficult to reduce to isolated motif labels. Archaeometric and conservation studies likewise show that lacquer objects are materially and technically heterogeneous, with craft practices varying across periods and object groups.¹⁻⁶ Museum catalogues can identify dragons, flowers, clouds, borders or landscape elements, but the visual organization of those elements also depends on role, adjacency, enclosure, centering, repetition and other relations. Comparable computational work has addressed object recognition, collection-scale visual analysis and machine-assisted study of museum or art-image datasets.⁷⁻¹⁰ A collection-scale description therefore needs to retain the object as the independent analytical unit while representing motifs and their observable spatial organization in a form that can be compared across a collection.

In this study, decorative grammar denotes a recurrent structure defined by motif identity, compositional role and canonical spatial relation across independent objects. Computational heritage has developed mature approaches to visual recognition, semantic interoperability and dataset documentation, but recurrent motif-role-spatial configurations have rarely been treated as auditable object-level units of analysis.⁷⁻¹⁴ The term is used descriptively: repeated configurations are not taken as evidence that historical makers followed a formal generative rule system, nor is symbolism inferred from co-occurrence alone. Separating object identity, observed relations and later interpretation is consistent with CIDOC CRM-oriented interoperability and evidence-aware heritage modelling.¹¹⁻¹⁴ The same separation also addresses established concerns about provenance, completeness and descriptive consistency in cultural-heritage data.¹⁵⁻²⁸

The analysis addresses four questions. Which motif classes, roles and spatial relations dominate the curated corpus? Which complete motif-role-relation configurations recur across independent objects?

How stable are motif and relation structures after excluding adjudicated objects, confidence-filtered records and negative controls? Finally, which coarse chronological differences remain when missing period metadata is treated as missingness rather than as a historical category and multiplicity is examined explicitly? The aim is not automated lacquerware classification, but a reproducible description of decorative configuration with clear evidential limits.

Results

Corpus structure

The final analytical corpus comprised 509 objects linked back to museum records and image evidence (Table 1). The frozen master contained 71 motif classes and 2,054 canonical relation triplets. Mean motif richness was 4.38 classes per object (median 4) and mean relation richness was 4.04 triplets per object (median 4). The corpus was dominated by Qing-period objects (374, 73.5%), followed by Ming objects (83, 16.3%); 31 objects had unresolved period metadata and 10 were grouped as Song-Yuan/Jin. The remaining 11 objects comprised Ming-Qing transition (5), Late Qing-Republic (4) and Republic/Modern (2) groups. The 56 adjudicated objects represented 11.0% of the analytical corpus. These smaller transition and modern groups remained in the master for traceability but were not used for primary chronological inference.

Table 1
Corpus and sample characteristics.

Characteristic	Value
Objects	509
Motif classes	71
Relation triplets	2,054
Mean motif richness	4.38
Median motif richness	4
Mean relation richness	4.04
Median relation richness	4
Adjudicated objects	56 (11.0%)
Period Qing	374
Period Ming	83
Period unresolved	31
Period Song-Yuan/Jin	10
Period Ming-Qing transition	5
Period Late Qing-Republic	4
Period Republic/Modern	2

Motif prevalence and compositional roles

FOLIAGE_SCROLL was the most prevalent motif, occurring on 290 objects (57.0%), followed by KEY_FRET (207, 40.7%), DIAPER (180, 35.4%), FLOWER_GENERIC (176, 34.6%), CLOUD (145, 28.5%) and DRAGON (112, 22.0%) (Table 2; Fig. 1). Role distributions further distinguished structural from pictorial usage: KEY_FRET was overwhelmingly BORDER, DIAPER was mainly STRUCTURAL or BORDER, CLOUD was primarily STRUCTURAL, DRAGON was primarily PRIMARY, and FLOWER_GENERIC was mainly SECONDARY (Table 3).

Table 2
Most prevalent motif classes at object level.

Motif	Objects	Prevalence
FOLIAGE_SCROLL	290	57.0%
KEY_FRET	207	40.7%
DIAPER	180	35.4%
FLOWER_GENERIC	176	34.6%
CLOUD	145	28.5%
DRAGON	112	22.0%
MOUNTAIN_ROCK	97	19.1%
FIGURE_GROUP	90	17.7%
PINE	77	15.1%
PAVILION	72	14.1%

Table 3
Motif-role distributions for the ten most prevalent motif classes. Role counts need not sum to the object count because unresolved role assignments were retained rather than forced, and a motif could rarely occupy more than one contextual role within an object.

Motif	Objects	Primary	Secondary	Structural	Border
FOLIAGE_SCROLL	290	6	42	205	34
KEY_FRET	207	0	0	1	202
DIAPER	180	1	0	126	45
FLOWER_GENERIC	176	47	115	0	6
CLOUD	145	5	26	112	0
DRAGON	112	91	19	0	0
MOUNTAIN_ROCK	97	0	29	67	0
FIGURE_GROUP	90	64	25	0	0
PINE	77	0	58	17	0
PAVILION	72	2	69	0	0

Co-occurrence and spatial relations

The most frequent motif co-occurrence was FLOWER_GENERIC with FOLIAGE_SCROLL (153 objects; Jaccard = 0.489). Other common pairs included FOLIAGE_SCROLL-KEY_FRET (107; Jaccard = 0.274), DIAPER-FOLIAGE_SCROLL (106; Jaccard = 0.291), and CLOUD-DRAGON (95; Jaccard = 0.586). FIGURE_GROUP-MOUNTAIN_ROCK (76; Jaccard = 0.685) and MOUNTAIN_ROCK-PINE (74; Jaccard = 0.740) showed particularly strong object-level association (Table 4; Fig. 2). These measures describe co-occurrence only and do not by themselves establish cultural or production causation.

Table 4
Selected frequent and high-association motif co-occurrence pairs and object-level Jaccard association.

Motif A	Motif B	Co-occurrence N	Jaccard
FLOWER_GENERIC	FOLIAGE_SCROLL	153	0.489
FOLIAGE_SCROLL	KEY_FRET	107	0.274
DIAPER	FOLIAGE_SCROLL	106	0.291
CLOUD	DRAGON	95	0.586
FIGURE_GROUP	MOUNTAIN_ROCK	76	0.685
MOUNTAIN_ROCK	PINE	74	0.740

Among the 2,054 canonical triplets, SURROUNDS was most common (492, 24.0%), followed by ADJACENT_TO (478, 23.3%), CENTER_OF (406, 19.8%) and BORDER_OF (360, 17.5%). INSIDE, REPEATED_AROUND, ABOVE, OVERLAPS and SYMMETRICAL_WITH accounted for the remainder (Table 5). The directed network showed repeated connections between motif classes and composition, panel, lid and body regions (Fig. 3), indicating that spatial organization is structured rather than reducible to motif frequency alone.

Table 5
Canonical spatial-relation type distribution.

Relation type	Triplets	Share
SURROUNDS	492	24.0%
ADJACENT_TO	478	23.3%
CENTER_OF	406	19.8%
BORDER_OF	360	17.5%
INSIDE	124	6.0%
REPEATED_AROUND	90	4.4%
ABOVE	54	2.6%
OVERLAPS	41	2.0%
SYMMETRICAL_WITH	9	0.4%

Recurrent role-contextualized configurations and relation robustness

To test the decorative-grammar claim directly, relations were summarized as normalized source motif class/role + relation code + target motif class or compositional region. Recurrent configurations were not inferred from marginal frequencies alone. The most common configuration was KEY_FRET[BORDER]-BORDER_OF->OBJECT_COMPOSITION (137 objects, 26.9%), followed by DRAGON[PRIMARY]-CENTER_OF->OBJECT_COMPOSITION (62, 12.2%) and CLOUD[STRUCTURAL]-SURROUNDS->DRAGON[PRIMARY] (48, 9.4%) (Table 6). A motif-to-motif-only summary, excluding composition and region targets, is provided in Table S4. Excluding all 56 adjudicated objects left 453 objects and 1,730 relation triplets; exact triplet weights remained correlated with the full corpus (Spearman rank correlation = 0.790), 19 of the top 20 triplets were retained, node weighted-degree rank correlation = 0.831, and relation-type rank correlation = 0.967. These sensitivity results show that the principal relation structure is not produced solely by adjudication, while the low pre-adjudication exact-triplet reliability still requires conservative interpretation.

Table 6

Top recurrent role-contextualized relation configurations and sensitivity to exclusion of adjudicated objects.

Configuration	Full n	Support	Non-adjudicated n	Non-adjudicated support
KEY_FRET[BORDER] - BORDER_OF -> OBJECT_COMPOSITION	137	26.9%	126	27.8%
DRAGON[PRIMARY] - CENTER_OF -> OBJECT_COMPOSITION	62	12.2%	61	13.5%
CLOUD[STRUCTURAL] - SURROUNDS -> DRAGON[PRIMARY]	48	9.4%	42	9.3%
FIGURE_GROUP[PRIMARY] - CENTER_OF -> OBJECT_COMPOSITION	43	8.4%	43	9.5%
DIAPER[BORDER] - BORDER_OF -> OBJECT_COMPOSITION	41	8.1%	40	8.8%
FIGURE_GROUP[PRIMARY] - ADJACENT_TO -> PAVILION[SECONDARY]	26	5.1%	26	5.7%
PEONY[PRIMARY] - CENTER_OF -> OBJECT_COMPOSITION	26	5.1%	25	5.5%
FOLIAGE_SCROLL[STRUCTURAL] - SURROUNDS -> OBJECT_COMPOSITION	26	5.1%	26	5.7%
FLOWER_GENERIC[PRIMARY] - CENTER_OF -> OBJECT_COMPOSITION	24	4.7%	23	5.1%
FOLIAGE_SCROLL[BORDER] - BORDER_OF -> OBJECT_COMPOSITION	22	4.3%	21	4.6%

Metadata-linked period patterns

Primary chronological inference was restricted to Ming (n = 83) and Qing (n = 374) objects. The Song-Yuan/Jin group (n = 10) and smaller historical groups were retained descriptively, while the unresolved group (n = 31) was explicitly treated as missing metadata rather than a historical period. Qing objects had mean motif richness 4.64 versus 3.93 for Ming objects and mean relation richness 4.19 versus 3.63. Permutation rank tests yielded $P = 0.0025$ for motif richness and $P = 0.0045$ for relation richness. The corresponding Qing-minus-Ming mean differences were 0.71 motifs (95% bootstrap CI 0.20 to 1.18; Cliff's $\delta = 0.218$) and 0.56 relations (95% bootstrap CI -0.03 to 1.09; Cliff's $\delta = 0.204$). The permutation rank tests and bootstrap intervals address different estimands: the former assess rank-distribution differences, whereas the latter summarize uncertainty in arithmetic mean differences; consequently, the relation-richness rank test can be significant while the mean-difference interval includes zero. For motif prevalence, FOLIAGE_SCROLL, KEY_FRET and DIAPER remained different after both Holm and

Benjamini-Hochberg sensitivity adjustment; no other top-ten motif survived either adjustment (Table 7; Fig. 4 and Fig. 5). This primary analysis therefore separates interpretable historical groups from metadata missingness and treats period effects as exploratory rather than as a continuous historical trajectory.

Table 7
Ming-Qing motif prevalence permutation tests with Holm family-wise and Benjamini-Hochberg false-discovery-rate sensitivity adjustment.

Motif	Ming %	Qing %	Permutation P	Holm-adjusted P	BH q
FOLIAGE_SCROLL	73.5	55.1	0.0015	0.012	0.005
KEY_FRET	19.3	45.7	0.0005	0.005	0.0025
DIAPER	15.7	42.2	0.0005	0.005	0.0025
FLOWER_GENERIC	31.3	35.3	0.539	1.000	0.674
CLOUD	27.7	29.7	0.786	1.000	0.786
DRAGON	21.7	23.5	0.775	1.000	0.786
MOUNTAIN_ROCK	13.3	21.1	0.1015	0.555	0.169
FIGURE_GROUP	15.7	19.8	0.4395	1.000	0.628
PINE	9.6	17.4	0.0925	0.555	0.169
PAVILION	8.4	17.1	0.0675	0.473	0.169

Reliability and sensitivity

The final database is an adjudicated analytical release, not an independently double-coded dataset meeting the pre-specified full-corpus reliability threshold. In the frozen 72-object pre-adjudication QC set, mean motif-set Jaccard was 0.653, mean relation-triplet Jaccard was 0.314, role Gwet AC1 was 0.697 and major disagreement affected 51/72 objects (Table S1 and Table S3). The 12 pre-specified negative-control objects were used to detect motif/relation false positives and protect the FORM-not-equal-MOTIF boundary; the control gate was closed, but the overall reliability gate remained HOLD because reliability and cross-batch drift criteria were not met. All targeted adjudications were conducted by a third adjudicator independent of Coder A and Coder B, using the frozen coding manual and the underlying multi-view image evidence. These metrics are retained without replacement by post-adjudication concordance.

Sensitivity analyses nevertheless showed that major collection-level motif rankings were stable. Excluding the 56 adjudicated objects gave Spearman rank correlation = 0.984 relative to the full corpus. Restricting the final analytical master to the 107 objects with Confidence_FINAL=HIGH gave Spearman rank correlation = 0.919, with a maximum absolute motif-prevalence difference of 6.54 percentage

points. Excluding the 12 negative controls (N = 497) produced Spearman rank correlation = 0.999975, preserved the top-10 motif set and changed any motif prevalence by at most 1.174 percentage points (Table S1, Table S2 and Fig. S1). Relation-level sensitivity is reported separately above because motif-rank stability cannot substitute for reliability of exact spatial triplets.

Discussion

The results support moving beyond motif inventories toward relation-aware descriptions of decoration. Table 6 shows recurrent complete motif-role-relation configurations rather than inferring grammar from marginal motif or relation frequencies. KEY_FRET[BORDER]-BORDER_OF->OBJECT_COMPOSITION and CLOUD[STRUCTURAL]-SURROUNDS->DRAGON[PRIMARY], for example, recur across independent objects. Network analysis provides a compact representation of these repeated links.^{29–32} Together, these observations provide the empirical basis for the operational decorative-grammar framework used here.

That term should not be confused with a historical claim that artisans consciously followed a formal generative rule system. The current data show recurrent visual organization in a museum corpus; they do not identify workshop practice, transmission pathways or symbolic meaning without additional historical evidence. This distinction is especially important for lacquerware, where material analyses, technical studies and historical interpretation document substantial variation in substrates, coatings, materials and production practices.^{1–6} The strong CLOUD-DRAGON and FIGURE_GROUP-MOUNTAIN_ROCK associations, for example, are valid descriptive findings, but their iconographic interpretation requires documentary or art-historical sources that are independent of the coding itself.

This evidential boundary is consistent with established art-historical scholarship on Chinese lacquer, which has treated ornament as historically structured decorative programs rather than as disconnected labels and has documented recurring repertoires that include abstract scrollwork, floral and bird subjects, figural landscapes, dragons, auspicious emblems and geometric or patterned grounds.^{33–35} The present workflow does not replace connoisseurship or iconographic interpretation; instead, it makes the observable spatial organization of such decorative repertoires explicit and auditable at collection scale.

The period-linked results are best treated as exploratory. Primary inference was limited to Ming and Qing objects, while unresolved metadata was removed from the chronological test and the small Song-Yuan/Jin group was interpreted descriptively. Qing objects showed higher average motif and relation richness, but effect sizes were modest and the bootstrap interval for the Qing-Ming relation-richness mean difference included zero. Only FOLIAGE_SCROLL, KEY_FRET and DIAPER retained motif-prevalence differences after Holm and Benjamini-Hochberg sensitivity adjustment. These patterns identify targets for historical follow-up rather than a continuous developmental trajectory. The National Palace Museum collection remains the empirical frame, and collection formation and digitization choices may influence what is visible in the data. More generally, dataset-bias research shows that apparently stable visual

patterns can partly reflect construction and selection of the source corpus rather than the full target domain.³⁶

The principal methodological limitation is the pre-adjudication reliability gate. Although early pilot and targeted stress tests supported structural freezing of the vocabulary, later production QC revealed cross-batch instability, particularly for exact motif and relation configurations. Reliability coefficients are sensitive to prevalence, category structure and the distinction between raw agreement and chance-corrected agreement; for that reason, no single coefficient was treated as sufficient evidence of coding quality.^{37–45} The relation-level sensitivity analysis is reassuring at the collection-structure level (19/20 top triplets retained after exclusion of adjudicated objects), but it does not erase the low pre-adjudication exact-triplet Jaccard. The final adjudicated master is therefore appropriate for transparent descriptive analysis only when the original reliability metrics, adjudication burden and sensitivity analyses are reported alongside it.

The main methodological contribution is therefore a representation strategy for complex decorated heritage: the object remains the unit of inference, motif occurrence is separated from compositional role, directional spatial relations are encoded explicitly, disagreement is preserved before adjudication, and collection-level patterns are tested against plausible exclusions. These choices are compatible with semantic-interoperability, provenance and dataset-documentation practices in digital cultural heritage and research-data stewardship.^{11–14,17–26} This structure allows recurrent decoration to be analysed without turning observable relationships into unsupported cultural interpretation.

Methods

Corpus and evidence unit

The analysis used 509 lacquerware objects linked to National Palace Museum, Taipei object records and corresponding image evidence. The object, rather than individual motif occurrences or images, was the independent unit for prevalence, co-occurrence, configuration support and metadata-linked comparisons. Object identifiers were blinded during production coding and later reconnected to the unblinded analytical master after coding and quality-control procedures were frozen. For the revised primary analysis, chronological inference was restricted to the Ming-Qing comparison; unresolved period metadata was treated as missingness rather than a historical category, and smaller period groups were retained for descriptive context only.

Motif occurrence, role and relation ontology

The observation model separated motif occurrence, motif identity, compositional role and spatial relation. Motif classes were encoded using a frozen controlled vocabulary. Roles used PRIMARY, SECONDARY, STRUCTURAL and BORDER in the final analytical summaries. Each relation was stored as SOURCE_OCCURRENCE-relation-TARGET and normalized for graph-like comparison to SOURCE motif class/role + relation code + TARGET motif class or region type; targets could be MOTIF, COMPOSITION or

REGION. Directional relations therefore preserve source-to-target meaning (for example CLOUD-SURROUNDS->DRAGON), whereas symmetric relations such as ADJACENT_TO and SYMMETRICAL_WITH were treated as non-directional associations for interpretation. This separation follows the general logic of relation-aware cultural-heritage semantic modelling while remaining deliberately narrower than a full ontology implementation.^{11–14} Form descriptors were not admitted as motifs solely because they described vessel geometry; ambiguous visual evidence could remain unidentified rather than being forced into a semantic class.

Independent coding, reliability and adjudication

Coding proceeded through independent blinded production and pre-specified QC subsets, followed by freeze, reliability calculation, targeted adjudication and canonical normalization. Early pilot and stress gates were used to stabilize observation rules before full production. The final 72-object QC set contained 40 CORE-RELATION, 20 SIMPLE-ELIGIBLE and 12 NEGATIVE-CONTROL objects; negative controls were designed to detect false-positive motif/relation coding and to enforce the FORM-not-equal-MOTIF boundary. Reliability was calculated before adjudication and retained as the definitive independent-coding evidence. Agreement was evaluated with set-overlap measures and chance-corrected coefficients rather than percent agreement alone, consistent with established guidance on inter-rater and inter-coder reliability.^{37–45} Because the full reliability gate remained HOLD, subsequent adjudication produced a canonical analytical release but did not retroactively convert the independent-coding gate to PASS.

Third-party adjudication protocol. Cases entered adjudication only after the independent coder records had been locked and the pre-adjudication reliability calculations had been preserved. Adjudication was performed by a third adjudicator who was independent of Coder A and Coder B. The adjudicator reviewed the frozen A/B records together with all available object views and visible surfaces, applying the already frozen motif, role and relation rules. The adjudicator could resolve disagreements within the existing vocabulary but could not introduce post hoc motif classes or relation codes to improve agreement. Original Coder A and Coder B records were retained unchanged for audit and reliability reporting; only the adjudicated decision was written to the canonical analytical master. Adjudication decisions were based on observable image evidence and the frozen coding rules, not on desired downstream statistical outcomes, and were completed before the final statistical analysis. Post-adjudication concordance was therefore treated as a production correction rather than as independent inter-coder reliability.

Statistical analysis

Motif prevalence counted each motif at most once per object. Pairwise co-occurrence was summarized by object count, Jaccard similarity and cosine association.^{46,47} Recurrent decorative configurations were counted at object level after combining normalized relation triplets with object-level role assignments; configuration support was the proportion of the 509 objects containing the configuration. No arbitrary inferential support threshold was imposed to classify configurations as grammatical or non-

grammatical; instead, recurrent configurations were ranked descriptively by object-level support, and decorative grammar refers to the observed repertoire of repeated motif-role-relation structures. Network figures were visualization summaries only: Fig. 2 was restricted to the 20 motifs with the highest weighted degree and Fig. 3 to the 18 motif or region nodes with the greatest triplet volume to maintain legibility; all statistical summaries and configuration analyses used the full analytical master. Primary period richness comparisons used the Kruskal-Wallis rank statistic.⁴⁸ Relation-level robustness compared the full master with the non-adjudicated subset using Spearman rank correlation,⁴⁹ top-20 triplet overlap, weighted-degree rank correlation and relation-type rank correlation. For exact-triplet weight correlation, the comparison universe was the union of all 718 unique canonical triplets observed in the full corpus; triplets absent after exclusion of adjudicated objects were assigned weight zero. Permutation procedures used 1,999 random label permutations, with permutation inference following standard resampling practice and non-zero Monte Carlo p-value calculation.^{50,51} Motif-prevalence comparisons used chi-square statistics under label permutation.⁵² Motif-specific period P values were reported raw and with Holm family-wise and Benjamini-Hochberg false-discovery-rate sensitivity adjustment.^{53,54} Effect magnitude was summarized with Cliff's δ ⁵⁵ and nonparametric bootstrap confidence intervals for arithmetic mean differences. Because rank tests and mean-difference intervals target different estimands, they were interpreted jointly rather than expected to produce identical null decisions.

Use of artificial intelligence tools

ChatGPT (OpenAI, GPT-5.6 Sol; accessed 29 September 2026) was used for drafting support, language editing, manuscript organization, code troubleshooting and quality-control scripting. The authors defined the research questions and coding rules, verified source materials and numerical results, checked the final analyses against the frozen analytical tables, reviewed all generated text and code, and take responsibility for the submitted work.

Data Availability

The unblinded object-level analytical master and derived analytical tables needed to audit and reproduce the reported results are provided as Supplementary Data 1. Stable object identifiers, period metadata, motif and role assignments, normalized relation triplets, confidence/adjudication flags and the fields required for the reported sensitivity analyses are retained in the analytical release. Museum image files are not redistributed in the supplementary package; object identifiers and derived annotations preserve traceability to the source catalogue without republishing third-party image assets. This approach is consistent with FAIR-oriented practice in which reusable derived data, provenance and access conditions are documented even when source assets cannot be redistributed.^{20,21,23}

Code Availability

A reproducibility package accompanies the submission and contains the frozen-value manifest, analysis notes, environment information and scripts for checking the principal summary tables, recurrent-

configuration counts and sensitivity statistics against Supplementary Data 1. The reproducibility checks do not require redistribution of the museum image files.

Declarations

Acknowledgements

The authors received no external funding for this work.

Author Contributions

P.Y. conceived the study, developed the methodology, performed the formal analysis and investigation, and contributed to validation. X.Z. developed software and contributed to validation and visualization. J.H. contributed to methodology and resources. P.Y., X.Z. and J.H. drafted, reviewed and edited the manuscript, and all authors approved the final version.

Competing Interests

The authors declare no financial or non-financial competing interests.

Ethics Statement

Ethics approval was not applicable. The study analysed publicly accessible cultural-heritage records, museum object metadata and image evidence and did not recruit human participants or animals or collect identifiable personal data.

References

1. Sung M, Jung J, Lu R, Miyakoshi T. Study of historical Chinese lacquer culture and technology - analysis of Chinese Qin-Han dynasty lacquerware. *J Cult Herit.* 2016;21:889–93. 10.1016/j.culher.2016.05.004.
2. Qin Y, et al. Exploring thitsi in Qing Dynasty lacquerware: insights from a preliminary study. *npj Herit Sci.* 2025;13:46. 10.1038/s40494-025-01568-3.
3. Schilling MR, Heginbotham A, van Keulen H, Szelewski M. Beyond the basics: A systematic approach for comprehensive analysis of organic materials in Asian lacquers. *Stud Conserv.* 2016;61(sup3):3–27. 10.1080/00393630.2016.1230978.
4. Fu Y, Chen Z, Zhou S, Wei S. Comparative study of the materials and lacquering techniques of the lacquer objects from Warring States Period China. *J Archaeol Sci.* 2020;114:105060. 10.1016/j.jas.2019.105060.
5. Wu M, Zhang B, Jiang L, Wu J, Sun G. Natural lacquer was used as a coating and an adhesive 8000 years ago, by early humans at Kuahuqiao, determined by ELISA. *J Archaeol Sci.* 2018;100:80–7. 10.1016/j.jas.2018.10.004.

6. Lu R, Miyakoshi T. *Lacquer Chemistry and Applications*. Amsterdam: Elsevier; 2015. 10.1016/C2014-0-04817-4.
7. Mensink T, van Gemert J. The Rijksmuseum Challenge: Museum-Centered Visual Recognition. In *Proceedings of International Conference on Multimedia Retrieval* 451–454ACM, Glasgow, (2014). 10.1145/2578726.2578791
8. Crowley EJ, Zisserman A. In Search of Art. *Computer Vision - ECCV 2014 Workshops. Lecture Notes in Computer Science*. Volume 8925. Cham: Springer; 2015. pp. 54–70. 10.1007/978-3-319-16178-5_4.
9. Fiorucci M, et al. Machine Learning for Cultural Heritage: A Survey. *Pattern Recognit Lett*. 2020;133:102–8. 10.1016/j.patrec.2020.02.017.
10. Cetinic E, She J. Understanding and Creating Art with AI: Review and Outlook. *ACM Trans Multimedia Comput Commun Appl* 18 Article. 2022;66:1–22. 10.1145/3475799.
11. Doerr M. The CIDOC Conceptual Reference Module: An ontological approach to semantic interoperability of metadata. *AI Mag*. 2003;24:75–92. 10.1609/aimag.v24i3.1720.
12. International Organization for Standardization. *ISO 21127:2023 Information and documentation - A reference ontology for the interchange of cultural heritage information*. Geneva: ISO; 2023.
13. Huo K, Mohd Noor NFB, Ismail RB. A CIDOC CRM-aligned Motif Configuration Application Profile for configurational meaning in Peranakan Nyonya ceramic heritage. *npj Herit Sci*. 2026. 10.1038/s40494-026-02954-1.
14. Hyvönen E. Using the Semantic Web in digital humanities: Shift from data publishing to data-analysis and serendipitous knowledge discovery. *Semantic Web*. 2020;11:187–93. 10.3233/SW-190386.
15. Candela G. *J Assoc Inf Sci Technol*. 2023;74:866–78. 10.1002/asi.24761. An automatic data quality approach to assess semantic data from cultural heritage institutions.
16. Lorenzini M, Rospocher M, Tonelli S. Automatically evaluating the quality of textual descriptions in cultural heritage records. *Int J Digit Libr*. 2021;22:217–31. 10.1007/s00799-021-00302-1.
17. Király P, Büchler M. Measuring completeness as metadata quality metric in Europeana. In *2018 IEEE International Conference on Big Data* 2711–2720 (IEEE, Seattle, 2018). 10.1109/BigData.2018.8622487
18. Haslhofer B, Isaac A. data.europeana.eu: The Europeana Linked Open Data Pilot. In *Proceedings of the International Conference on Dublin Core and Metadata Applications* 94–104DCMI, (2011). 10.23106/dcmi.952135673
19. Isaac A, Haslhofer B. *Semantic Web*. 2013;4:291–7. 10.3233/SW-120092. Europeana Linked Open Data - data.europeana.eu.
20. Wilkinson MD, et al. The FAIR Guiding Principles for scientific data management and stewardship. *Sci Data*. 2016;3:160018. 10.1038/sdata.2016.18.
21. Gebru T, et al. Datasheets for Datasets. *Commun ACM*. 2021;64:86–92. 10.1145/3458723.

22. Bender EM, Friedman B. Trans Assoc Comput Linguist. 2018;6:587–604. 10.1162/tacl_a_00041. Data Statements for Natural Language Processing: Toward mitigating system bias and enabling better science.
23. Lee BCG. The Collections as ML Data checklist for machine learning and cultural heritage. J Assoc Inf Sci Technol. 2025;76:375–96. 10.1002/asi.24765.
24. Sambasivan N et al. ACM,. Everyone wants to do the model work, not the data work: Data Cascades in High-Stakes AI. In Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems, Article 39, 1–15 (2021). 10.1145/3411764.3445518
25. Mitchell M et al. ACM, Atlanta,. Model Cards for Model Reporting. In Proceedings of the Conference on Fairness, Accountability, and Transparency 220–229 (2019). 10.1145/3287560.3287596
26. Arnold M, et al. FactSheets: Increasing trust in AI services through supplier's declarations of conformity. IBM J Res Dev. 2019;63(1–6):13. 10.1147/JRD.2019.2942288.
27. Paolanti M, Frontoni E, Pierdicca R. Towards trustworthy AI in cultural heritage. npj Herit Sci. 2026;14:131. 10.1038/s40494-026-02403-z.
28. Colace F, Gaeta R, Lorusso A, Pellegrino M, Santaniello D. New AI challenges for cultural heritage protection: A general overview. J Cult Herit. 2025;75:168–93. 10.1016/j.culher.2025.07.019.
29. Newman MEJ. The structure and function of complex networks. SIAM Rev. 2003;45:167–256. 10.1137/S003614450342480.
30. Newman MEJ. Networks: An Introduction. Oxford: Oxford University Press; 2010. 10.1093/acprof:oso/9780199206650.001.0001.
31. Borgatti SP, Mehra A, Brass DJ, Labianca G. Network analysis in the social sciences. Science. 2009;323:892–5. 10.1126/science.1165821.
32. Wasserman S, Faust K. Social Network Analysis: Methods and Applications. Cambridge: Cambridge University Press; 1994. 10.1017/CBO9780511815478.
33. Garner HM. Chinese Lacquer. London: Faber and Faber; 1979.
34. Watt JCY, Ford BB. East Asian Lacquer: The Florence and Herbert Irving Collection. New York: Metropolitan Museum of Art; 1991.
35. Clifford D. Chinese Carved Lacquer. London: Bamboo Publishing; 1992.
36. Torralba A, Efros AA. IEEE, Colorado Springs,. Unbiased look at dataset bias. In 2011 IEEE Conference on Computer Vision and Pattern Recognition 1521–1528 (2011). 10.1109/CVPR.2011.5995347
37. Cohen J. A coefficient of agreement for nominal scales. Educ Psychol Meas. 1960;20:37–46. 10.1177/001316446002000104.
38. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics. 1977;33:159–74. 10.2307/2529310.
39. Feinstein AR, Cicchetti DV. High agreement but low kappa: I. The problems of two paradoxes. J Clin Epidemiol. 1990;43:543–9. 10.1016/0895-4356(90)90158-L.

40. Gwet KL. Computing inter-rater reliability and its variance in the presence of high agreement. *Br J Math Stat Psychol*. 2008;61:29–48. 10.1348/000711006X126600.
41. Gwet KL. *Handbook of Inter-Rater Reliability*. 4th ed. Gaithersburg, MD: Advanced Analytics; 2014.
42. Krippendorff K. *Content Analysis: An Introduction to Its Methodology*. 4th ed. Thousand Oaks, CA: SAGE; 2018.
43. Artstein R, Poesio M. Inter-Coder Agreement for Computational Linguistics. *Comput Linguist*. 2008;34:555–96. 10.1162/coli.07-034-R2.
44. Sim J, Wright CC. The kappa statistic in reliability studies: use, interpretation, and sample size requirements. *Phys Ther*. 2005;85:257–68. 10.1093/ptj/85.3.257.
45. McHugh ML. Interrater reliability: the kappa statistic. *Biochem Med*. 2012;22:276–82. 10.11613/BM.2012.031.
46. Kosub S. A note on the triangle inequality for the Jaccard distance. *Pattern Recognit Lett*. 2019;120:36–8. 10.1016/j.patrec.2018.12.007.
47. Salton G, McGill MJ. *Introduction to Modern Information Retrieval*. New York: McGraw-Hill; 1983.
48. Kruskal WH, Wallis WA. Use of ranks in one-criterion variance analysis. *J Am Stat Assoc*. 1952;47:583–621. 10.1080/01621459.1952.10483441.
49. Spearman C. The proof and measurement of association between two things. *Am J Psychol*. 1904;15:72–101. 10.2307/1412159.
50. Good P. *Permutation. Parametric and Bootstrap Tests of Hypotheses*. 3rd ed. New York: Springer; 2005. 10.1007/b138696.
51. Phipson B, Smyth GK. Permutation P-values should never be zero: calculating exact P-values when permutations are randomly drawn. *Stat Appl Genet Mol Biol*. 2010;9., Article 39 (10.2202/1544-6115.1585).
52. Agresti A. *Categorical Data Analysis*. 3rd ed. Hoboken, NJ: Wiley; 2013.
53. Holm S. A simple sequentially rejective multiple test procedure. *Scand J Stat*. 1979;6:65–70. 10.2307/4615733.
54. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc B*. 1995;57:289–300. 10.1111/j.2517-6161.1995.tb02031.x.
55. Cliff N. Dominance statistics: ordinal analyses to answer ordinal questions. *Psychol Bull*. 1993;114:494–509. 10.1037/0033-2909.114.3.494.

Figures

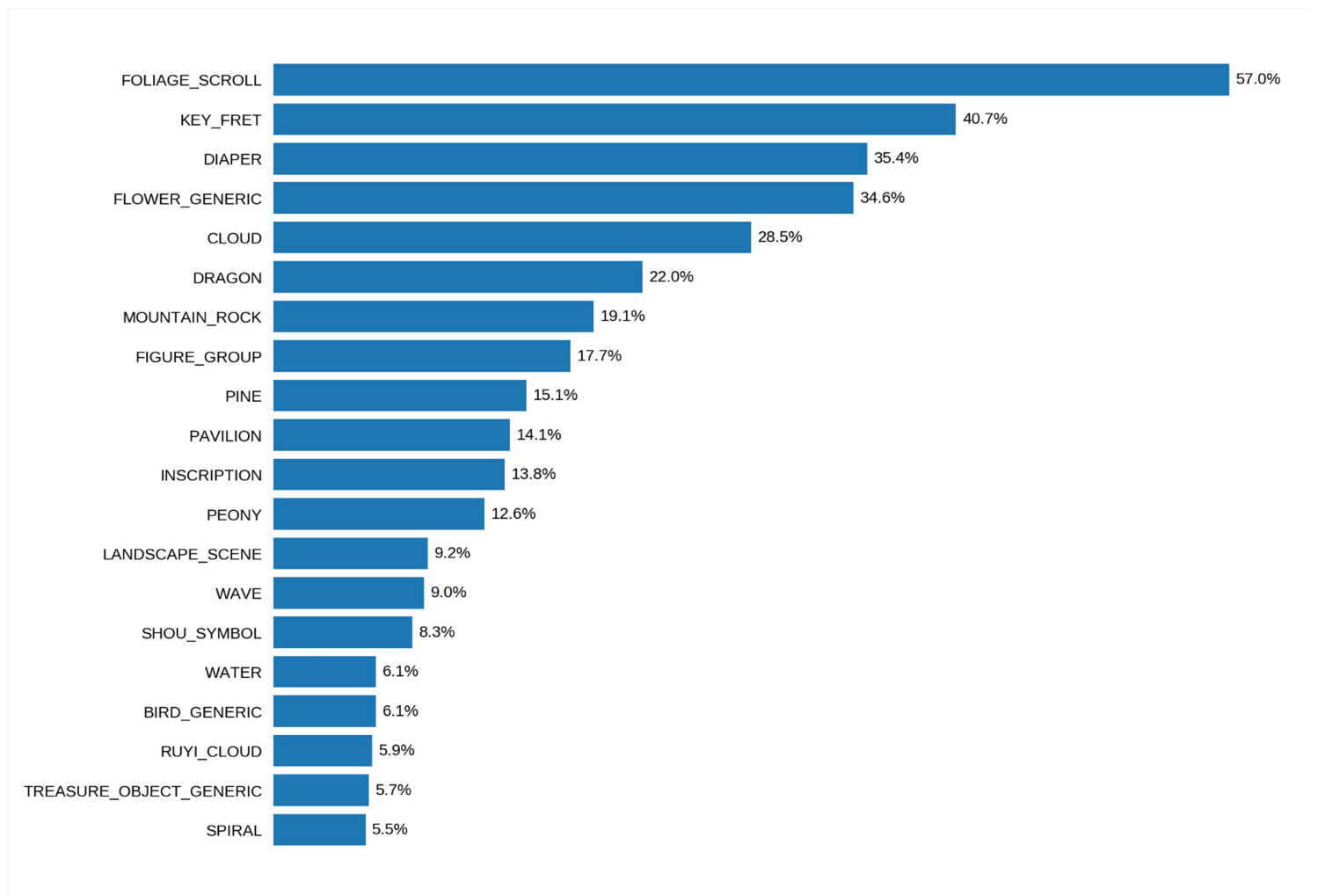


Figure 1

Object prevalence of the 20 most frequent motif classes. Each motif contributes at most once per object; denominator N=509.

Motif co-occurrence network (top 20 weighted-degree nodes)

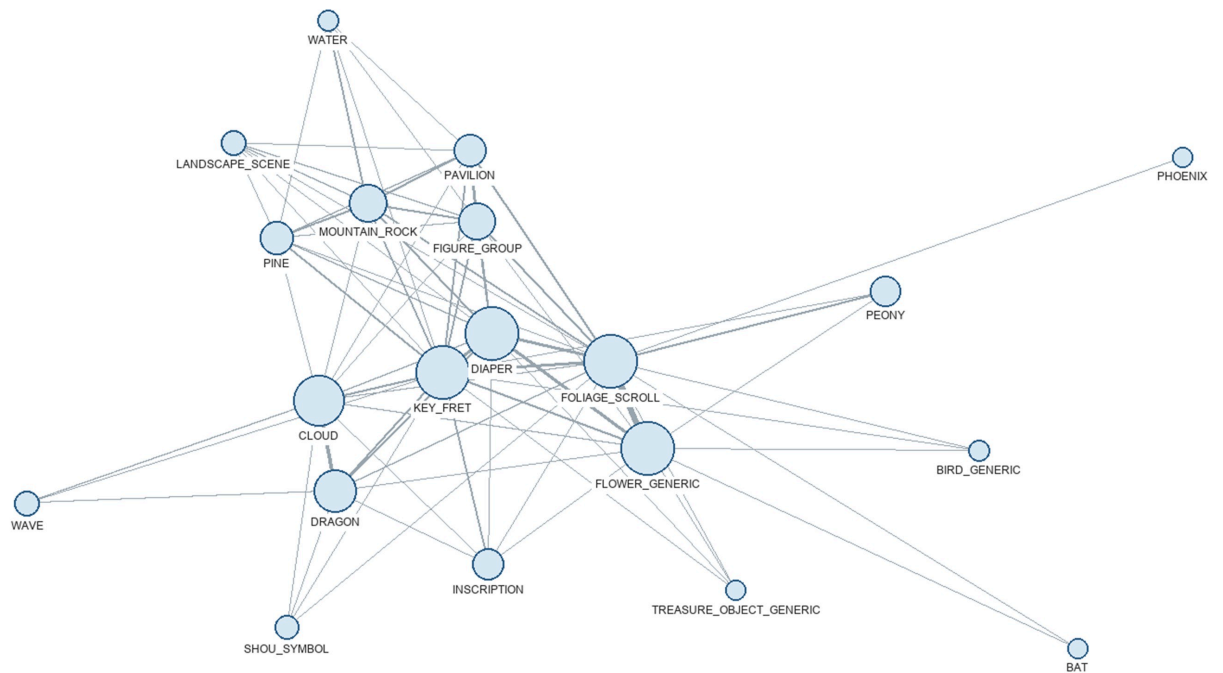


Figure 2

Motif co-occurrence network for the 20 motifs with the highest weighted degree. Edge weight is the number of objects in which the two motifs co-occur.

Spatial relation network of the 18 highest-volume motif and region nodes

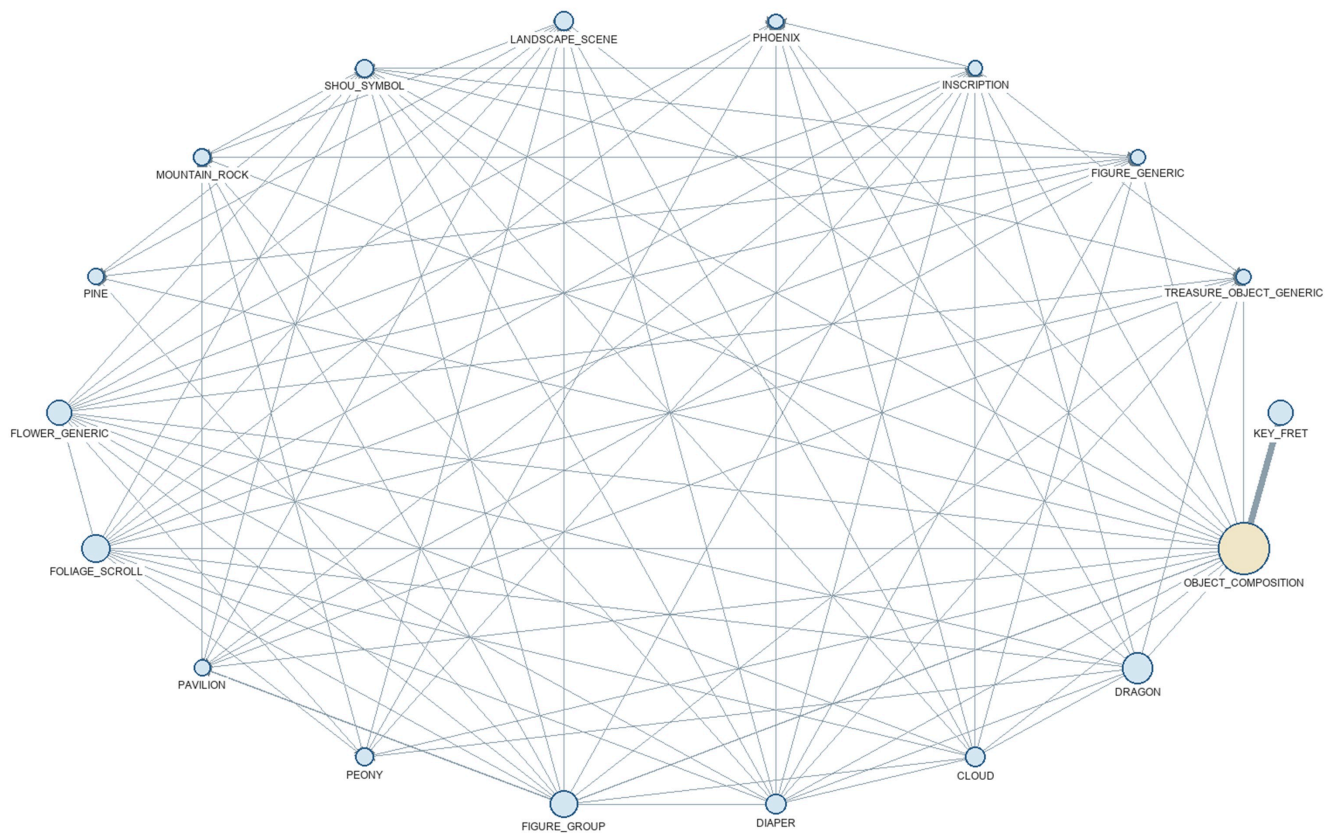


Figure 3

Directed spatial-relation network for the 18 motif or region nodes with the greatest triplet volume. Arrows follow canonical SOURCE-to-TARGET direction.

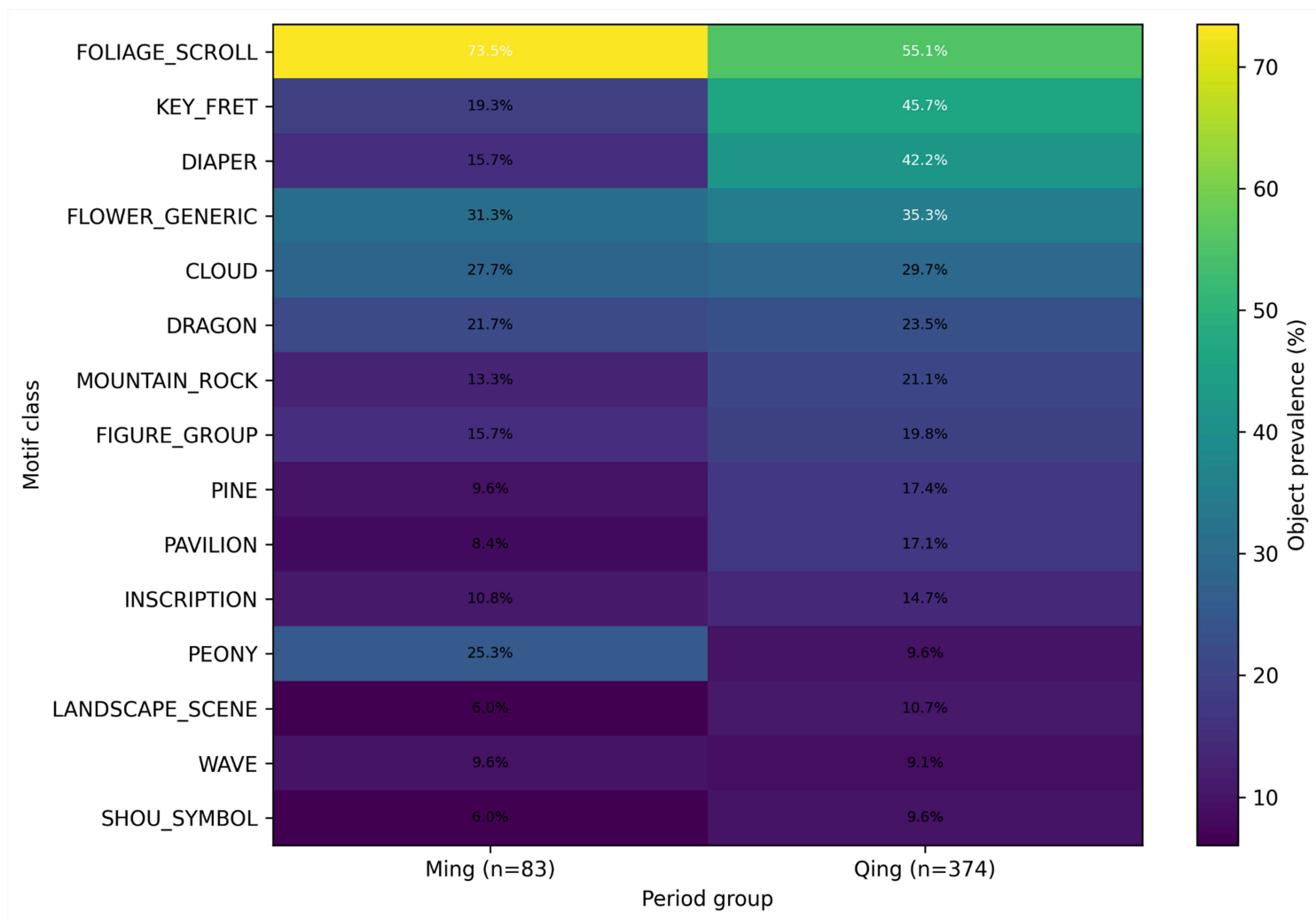


Figure 4

Object prevalence of the 15 most frequent motif classes in the primary Ming-Qing comparison. Percentages use the number of objects in each period group as denominator; unresolved and small-period groups are not used for primary inference.

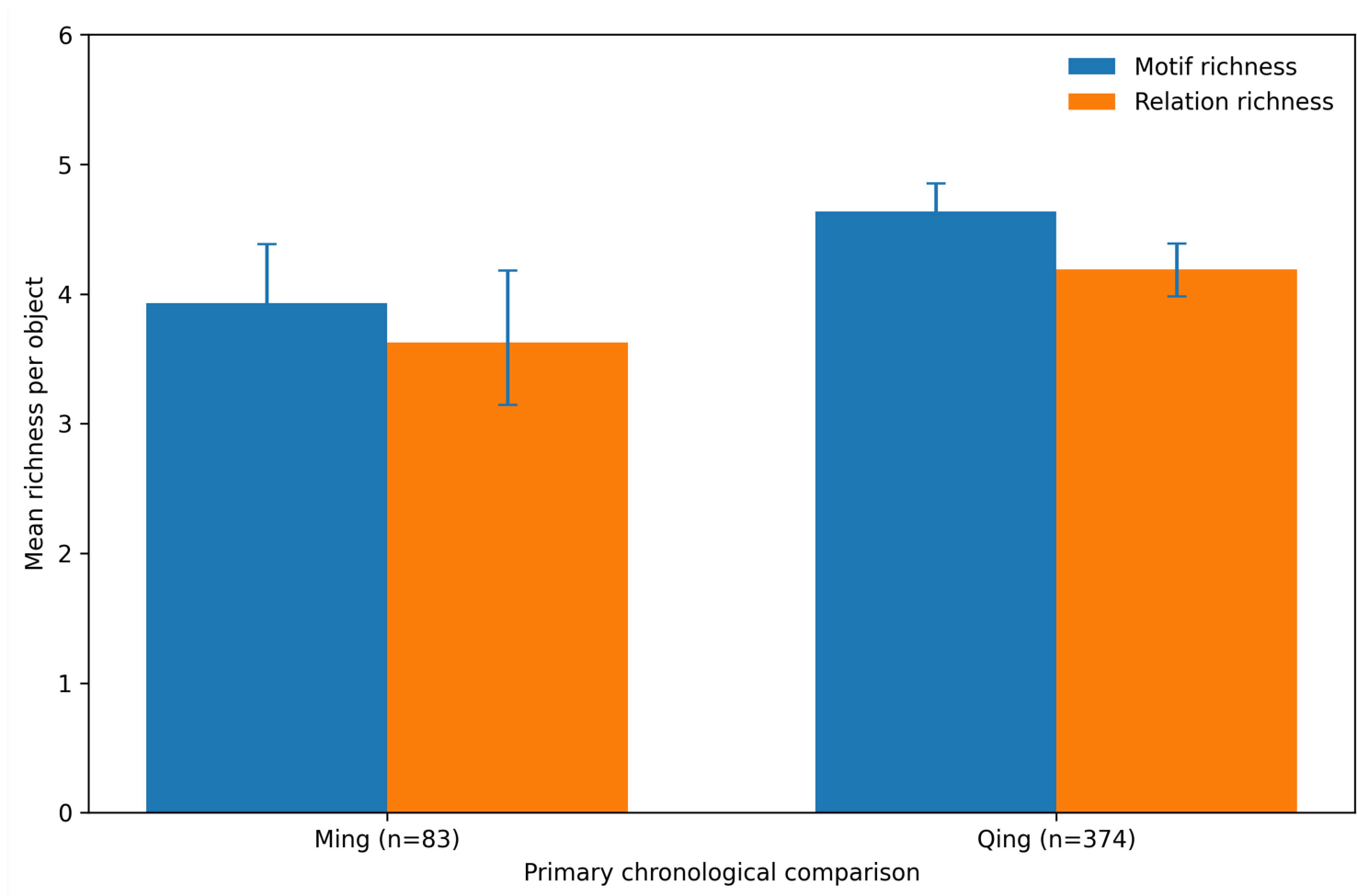


Figure 5

Mean motif and relation richness for Ming and Qing objects. Error bars show nonparametric bootstrap 95% confidence intervals for the group means.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [P08CReproducibilityCodeandREADMEFINAL.zip](#)
- [P08CSupplementaryData1FINAL.xlsx](#)
- [P08CSupplementaryInformationFINAL.pdf](#)