

Menu Genomics: Core Vocabulary, Phylogeny, and the Combinatorial Frontier of a Quick-Service Kitchen

Dave Mattson David Moss

Independent Research, Seattle, Washington

Computational analysis assisted by Claude (Anthropic)

Manuscript prepared in journal format, 8 September 2026. Not yet submitted or peer reviewed.

Independent study; not affiliated with or endorsed by Taco Bell Corp. or Yum! Brands, Inc. See Disclaimer.

Abstract. Quick-service restaurant menus are widely believed to be permutations of a small number of ingredients, but the claim is rarely quantified. We decompose all 76 food items on the September 2026 United States Taco Bell menu into 57 kitchen components (14 vehicles and 43 fillings) and treat the resulting presence-absence matrix as a collection of genomes. Eight fillings account for 55 % of all filling slots, twelve for 67 %, and a 25-component core (components used in five or more items) fills 82 % of all item-component slots; 38 % of items are pure recombinations of that core and a further 36 % differ from it by exactly one signature component, almost always the vehicle. Two thirds of items lie within two component edits of another item. Rarefaction of the ingredient vocabulary yields a Heaps exponent $\alpha \approx 1.06$, indicating a saturated (closed) pangenome in which new items are recombinations rather than additions. Average-linkage clustering on Jaccard distance (cophenetic correlation 0.81) and Louvain community detection on the item graph both recover three flavor systems—*Melt* (tortilla and melted three-cheese blend), *Fresh* (lettuce, cheddar, seasoned beef) and *Loaded* (nacho cheese sauce and sour cream over rice or beans)—that cut across the operator’s own nine menu categories (adjusted Rand index 0.23–0.28). Percolation analysis locates a sharp transition in menu connectivity between Jaccard 0.40 and 0.45 and identifies a small set of bridge items and bridge ingredients that hold the menu together. Pairwise linkage analysis reveals conserved motifs and, more informatively, 107 pairs of common ingredients that never co-occur, a grammar of mutually exclusive alleles (one protein, one tomato form, one cool cream) that the kitchen enforces implicitly. Link prediction on the ingredient co-occurrence graph and exhaustive enumeration of 1 205 088 grammar-valid candidate items map a familiarity–novelty landscape from which we propose twelve new products, from near-menu extensions to deliberately provocative boundary crossings, each scored against the existing menu.

Keywords: culinary combinatorics; pangenome; bipartite networks; community detection; link prediction; menu engineering; food product design.

1 Introduction

A recurring piece of folk wisdom about quick-service Mexican-inspired restaurants holds that the entire menu is “the same eight ingredients in different shapes.” The observation is intuitive, widely repeated, and, to our knowledge, has never been tested with any rigor. It is also more interesting than it first appears. If a kitchen with a fixed pantry generates dozens of distinct products by recombination, then the menu is not merely a list; it is a combinatorial system with a vocabulary, a grammar, and an unexplored space of well-formed but unproduced items. Those are exactly the properties that population genomics and network science were built to study.

This paper treats the Taco Bell United States menu as a corpus of genomes. Each menu item is a genome; each kitchen component (a tortilla, a protein, a sauce) is a gene; the presence-absence matrix of items by components is the object of study. The analogy is not decorative. Pangenome analysis asks how many distinct genes a population carries and whether that number is still growing [Tettelin et al., 2005, 2008, Medini et al., 2005]; the same question, asked of a menu, is whether new products bring new ingredients. Phylogenetic clustering asks which genomes descend from common ancestors; a menu

that grows by copy-and-modify has a literal lineage structure. Linkage disequilibrium asks which genes travel together and which are never found in the same genome; a kitchen has operational rules that produce precisely those signatures. And the bipartite item-ingredient graph, projected onto either side, supports the standard toolkit of centrality, community detection and link prediction [Newman, 2010, Blondel et al., 2008, Liben-Nowell and Kleinberg, 2007].

We make four contributions. First, we produce a clean, component-level decomposition of a complete quick-service menu at a single point in time and quantify how much of it is explained by how few components. Second, we show that the menu’s ingredient vocabulary is saturated in the pangenome sense and that its true organization is three flavor systems rather than the operator’s product categories. Third, we characterize the implicit grammar of the kitchen as a set of mutually exclusive ingredient alleles and conserved motifs. Fourth, we use link prediction and exhaustive enumeration to map the space of unbuilt items and propose twelve concrete products, each placed on a familiarity–novelty landscape and given a name.

Nothing in what follows required proprietary data. Every input is the public description of a menu item on the operator’s

ordering website.

2 Related work

The closest antecedent is the flavor-network programme of Ahn et al. [2011], who built a bipartite graph of recipes and ingredients across cuisines and projected it onto ingredients weighted by shared flavor compounds. Our projection is weighted by shared *items*, not compounds, and our corpus is a single commercial menu rather than a cross-cultural recipe collection, but the graph machinery is the same. Teng et al. [2012] used ingredient co-occurrence and substitution networks mined from recipe websites to predict recipe ratings; their substitution network is a close cousin of the allele structure we report in Section 5.6.

Pangenome analysis originated with Tettelin et al. [2005], who observed that the set of distinct genes in *Streptococcus agalactiae* kept growing as more strains were sequenced, and formalized the open/closed distinction through a power-law fit to the number of new genes per added genome [Tettelin et al., 2008]. The exponent of that fit is a Heaps exponent in the sense of Heaps [1978]; we adopt the Tettelin convention that $\alpha > 1$ indicates a closed pangenome.

In network science we rely on modularity-based community detection [Newman, 2006, Blondel et al., 2008], betweenness centrality [Freeman, 1977], percolation on thresholded similarity graphs, and neighborhood-based link prediction: the Adamic–Adar index [Adamic and Adar, 2003], the resource-allocation index [Zhou et al., 2009], and the Jaccard coefficient [Jaccard, 1901]. Agreement between partitions is measured by the adjusted Rand index [Hubert and Arabie, 1985]. Hierarchical clustering uses the unweighted pair-group method with arithmetic mean [Sokal and Michener, 1958]. Menu engineering as a business discipline [Kasavana and Smith, 1982] is concerned with the profitability and popularity matrix of items; it is orthogonal to the structural questions asked here, though our results bear on it (Section 7).

3 Data

3.1 Corpus

All data were collected from the operator’s public ordering website (tacobell.com) on 8 September 2026. We enumerated every category page (Tacos, Burritos, Specialties, Quesadillas, Nachos, Bowls, Sides & Sweets, Breakfast, Cantina Chicken, Luxe Value Menu, Veggie Cravings) and every product page reachable from them. Product pages expose a “What’s Included” list for most items; for ten items only a marketing description was available and the components were transcribed from it. Prices and calorie figures on the static pages are store-dependent placeholders and were discarded.

The corpus comprises 76 food items. Excluded were combination meals and boxes (bundles of items already counted), beverages, and standalone sauce packets. Where an item offers a protein choice, the default protein as listed was used;

the chicken and steak variants of beef items would add permutations but no new components.

3.2 Component taxonomy

Component names were normalized to 57 canonical kitchen components and assigned to one of eight roles: vehicle (14 components: flour tortilla, crunchy taco shell, Doritos shell, Tajín shell, white-corn shell, chalupa shell, gordita flatbread, tostada shell, Mexican Pizza shell, crispy tortilla bowl, tortilla chips, Nacho Fries, crispy chicken shell, stacker flatbread), protein (9), beans/starch (5), cheese (4), sauce/cream (10), produce (7), crunch/seasoning (3) and dessert (5). Normalization decisions with material effect on the results are: shredded cheddar and the three-cheese blend are kept distinct, because several items carry both; seasoned potatoes and “crispy potato bites” are one component; the crunchy, Doritos, Tajín and white-corn shells are distinct vehicles because they are distinct stock-keeping units; and the Cantina line’s two side salsa packets are counted as components because they ship with the item by default. The Tajín Taco’s “ranch sauce” is unnamed on the site and was mapped to Spicy Ranch to match its Cheesy Gordita Crunch sibling.

The resulting matrix $X \in \{0, 1\}^{76 \times 57}$ has 402 nonzero entries (item–component slots), of which 333 are filling slots and 69 are vehicle slots. Two items flagged “not currently available” on their pages (Quesarito, Chili Cheese Nacho Fries) were retained because their compositions are published.

4 Methods

4.1 Coverage and permutation structure

Component frequency is the column sum of X . The coverage curve adds fillings in descending frequency and records, at each step, the share of filling slots covered and the share of items whose fillings are entirely contained in the set (any vehicle permitted). A *core* was defined as components used in at least five items. Items were classified as pure permutations of the core, core plus one signature component, or core plus two or more. Edit distance between items is the size of the symmetric difference of their component sets. Families are connected components of the graph joining items at edit distance ≤ 2 that also share at least two components. A derivation forest was built by processing items in order of size and attaching each to the nearest already-placed item reachable by one of a small set of named operators (protein swap, shell swap, Supreme, add sauce, add protein, wrap, cheese swap).

4.2 Pangenome rarefaction

Items were sampled without replacement in random order and the number of distinct components observed after n items recorded, averaged over 300 permutations. Following Tettelin et al. [2008], the mean number of components new at step n was fitted by least squares on log–log axes to $\kappa n^{-\alpha}$ for $n \geq 2$. Components were also binned in the pangenome vocabulary:

core ($\geq 50\%$ of items), shell (10–50%), cloud (2–9%) and singleton.

4.3 Similarity, phylogeny and communities

Item similarity is the Jaccard index of component sets. UPGMA clustering was performed on Jaccard distance and validated by the cophenetic correlation coefficient; clades were cut at distance 0.75. A weighted item graph with all positive-Jaccard edges was partitioned by the Louvain method [Blondel et al., 2008] at resolution 1.0 (seed 7), and agreement with the operator’s categories was measured by the adjusted Rand index and normalized mutual information. Percolation was assessed by thresholding the item graph at Jaccard $\theta \in \{0.05, \dots, 1.00\}$ and recording the size of the largest connected component and the number of isolates.

4.4 Ingredient network

The bipartite item–component graph B was projected onto components, with edge weight equal to the number of shared items. Betweenness centrality was computed on B itself (so that a component’s score reflects how often it lies on shortest paths between items) and on the projection; eigenvector centrality was computed on the giant component of the projection, since desserts and one side are disconnected. Category spread of each component was summarized by the Shannon entropy of its distribution over menu categories.

4.5 Linkage and alleles

For every pair of components each present in three or more items, lift was computed as observed co-occurrence divided by the count expected under independence, f_{ab}/N , and significance by a two-sided Fisher exact test on the 2×2 contingency table. Pairs with zero co-occurrence, both components in five or more items, were reported as mutually exclusive. Conserved motifs were extracted as maximal sets of fillings (vehicles ignored) recurring in four or more items.

4.6 Link prediction and candidate enumeration

For every pair of components that never co-occur (excluding desserts and components in fewer than two items), the resource-allocation index [Zhou et al., 2009], Adamic–Adar index [Adamic and Adar, 2003] and Jaccard coefficient were computed on the unweighted projection. Same-role pairs were removed, as these are substitutes rather than opportunities.

Separately, candidate items were enumerated from a grammar over the existing components: one vehicle (or one of six observed vehicle pairs, or none for a bowl), one protein slot, one cheese configuration, one cool-cream configuration, one produce set, and one optional extra, subject to the constraints that a candidate has 3–9 components, is not an existing item, has a protein or a cheese, and, if a bowl, has a base starch. The daypart rule (eggs never with lettuce or sour cream) was

enforced in the enumeration and deliberately broken in one hand-built proposal. Each candidate was scored for *familiarity*, the mean Jaccard similarity to its three nearest existing items, and *novelty*, the number of its component pairs that never co-occur anywhere on the menu.

All analyses were performed in Python with NetworkX 3.6 and SciPy 1.17; code and the component matrix are available on request.

5 Results

5.1 A small core explains most of the menu

Figure 2 shows the coverage curve. The eight most frequent components—flour tortilla (39 items), three-cheese blend (32), cheddar (25), reduced-fat sour cream (25), seasoned beef (25), lettuce (24), nacho cheese sauce (21) and tomatoes (20)—fill 55.3% of filling slots once the tortilla is set aside as a vehicle; twelve fillings fill 67.0%, sixteen 75.1% and twenty 81.4%. Under the five-item core definition, 25 components fill 330 of 402 slots (82.1%). Twenty-nine items (38%) are built entirely from the core, 27 (36%) are core plus one signature component, and 20 (26%) carry two or more non-core components. Among the 27 single-signature items, the signature is a vehicle in 15 cases (crunchy shell, Doritos shell, chalupa shell, tostada, tortilla bowl, crispy chicken shell). The folk claim is therefore accurate in a specific form: fillings are shared, vehicles differentiate.

5.2 Permutation structure

Fifteen items (20%) are one edit from another item; 51 (67%) are within two edits; 63 (83%) within three. Under the family definition, 13 multi-item families absorb 43 items. The largest, with 13 members, chains from the Meximelt through the Cheesy Roll Up into every quesadilla and breakfast burrito, because tortilla plus three-cheese blend is the hub from which the flour-tortilla side of the menu radiates. The derivation forest attaches 29 items (38%) to a parent by a single named operator; the operators used most are protein swap (8 items; beef $\leftrightarrow$ black beans, bacon $\leftrightarrow$ sausage $\leftrightarrow$ steak), shell swap (4; crunchy $\leftrightarrow$ soft $\leftrightarrow$ Doritos $\leftrightarrow$ fries), add-a-sauce (4), Supreme (3; add sour cream and tomato) and add-a-protein (3). Three fillings—seasoned beef, lettuce, cheddar—appear on three different shells as three different named products, and their Supreme variants on the same three shells make six.

The items farthest from any other (edit distance five or six) are Chicken Enchilada Nacho Fries, Fresca Pepper Jack Steak Burrito, Naked Chicken Chalupa, Mini Taco Salad, Cantina Chicken Crispy Taco and the Avocado Ranch Chicken Stacker. Four of the six achieve their distinctiveness by placing ordinary fillings on an unusual vehicle.

5.3 The pangenome is closed

Figure 4 shows the rarefaction curves. Half the menu (38 items) already exposes 48.2 of the 57 components on average;

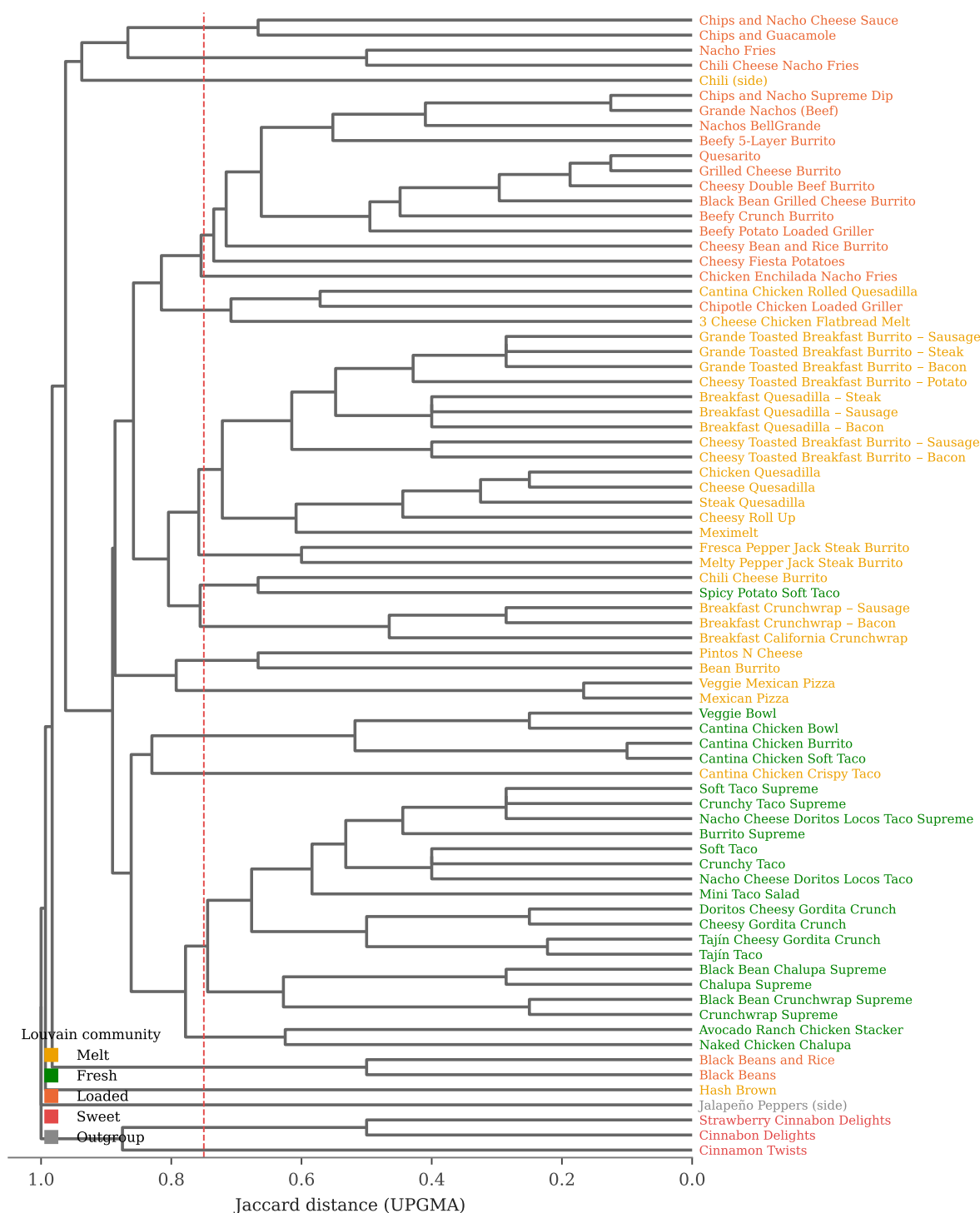


Figure 1: UPGMA dendrogram of all 76 items on Jaccard distance. Leaf labels are colored by Louvain community; the dashed line is the 0.75 cut yielding 21 clades. Cophenetic correlation 0.813.

the expected number of new components contributed by the 76th item is 0.11. The Tettelin fit gives $\alpha = 1.06$ ($\kappa = 16.7$). In the pangenome vocabulary the menu has one core gene (the flour tortilla is the only component in at least half of all items), 12 shell genes, 35 cloud genes and 9 singletons. With 76 genomes the exponent is not precise—a fit to the cumulative

curve rather than the increment curve gives a materially lower value—so we read the result qualitatively: the vocabulary is saturating, and new items are recombinations. This is consistent with the operator’s announced 2026 pipeline, whose items (Cheesy G Sliders, Mexican Pizza Empanadas, Firecracker Burrito, Queso Cracked Cantina Bowl) are new vehicles or

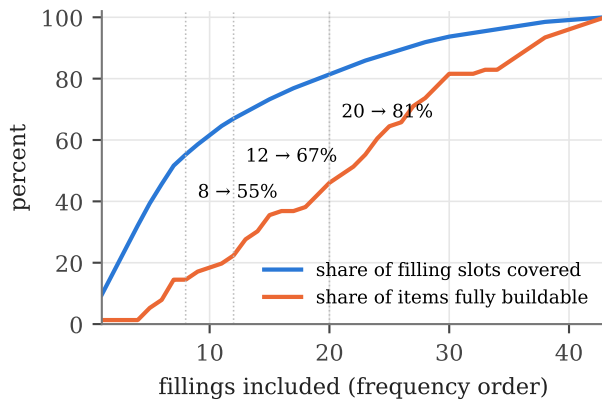


Figure 2: Coverage by number of core fillings, added in frequency order. Slot share rises quickly; the share of complete items lags because most items carry one signature component, usually a vehicle.

new formats around existing fillings.

5.4 Three flavor systems

UPGMA clustering on Jaccard distance has cophenetic correlation 0.813; the tree is shown in Figure 1. Cutting at distance 0.75 yields 21 clades. Louvain community detection on the weighted item graph finds five communities with modularity 0.206: three substantive systems of 29, 23 and 20 items, a three-item dessert community, and the isolated jalapeño side. We name the systems by their most enriched components:

Melt (29 items). Three-cheese blend (19 of 29, enrichment $1.6\times$) and flour tortilla ($1.5\times$). Every quesadilla, every breakfast item, the Meximelt, both Mexican Pizzas, the Cheesy Roll Up, the pepper-jack steak burritos and the Chili Cheese Burrito. Proteins vary freely because the melted cheese carries the flavor.

Fresh (23 items). Lettuce (23 of 23, $3.2\times$), cheddar ($2.4\times$), seasoned beef ($1.9\times$). Every crunchy-shell taco, the chalupas and gorditas, both Crunchwraps, the Mini Taco Salad, the Naked Chicken Chalupa, the Stacker, and the entire Cantina Chicken line, which clusters here through lettuce, cabbage and cheddar rather than through its chicken.

Loaded (20 items). Nacho cheese sauce (16 of 20, $2.9\times$) and reduced-fat sour cream ($2.0\times$), typically over rice or refried beans. Grilled Cheese Burrito, Quesarito, Beefy 5-Layer, Beefy Crunch, both Loaded Grillers, all nachos and chips, all fries, the fiesta potatoes.

The operator’s nine categories agree only weakly with either partition: adjusted Rand index 0.234 for the Louvain communities and 0.284 for the UPGMA clades, normalized mutual information 0.456. The category “Specialties” is split ten to five to two across the three systems; “Burritos” nine to four to two. The tree also produces individually informative placements. Burrito Supreme clusters with the Supreme tacos, not with other burritos; the Spicy Potato Soft Taco is filed

among the breakfast Crunchwraps (potato, cheddar, tortilla); the Cantina line forms its own clade adjacent to the tacos.

5.5 Hub ingredients and bridge items

The component projection has 384 of 1,596 possible edges (density 24.1 %); Figure 3 shows the subgraph of components in three or more items. Betweenness on the bipartite graph ranks the flour tortilla (0.184), three-cheese blend (0.165), cheddar (0.104), lettuce (0.072), seasoned beef (0.067), sour cream (0.061), tomatoes (0.057) and nacho cheese sauce (0.055) as the components most often on shortest paths between items; these are the octet. Chili is notable below them: present in only three items, it carries betweenness 0.015 because it alone links the sides shelf, the burrito lineage and the fries lineage. Category entropy identifies sour cream ($H = 2.39$ bits, seven categories) and the three-cheese blend (2.37) as the most promiscuous components, and confirms an asymmetric daypart barrier: of the breakfast genome’s 14 components, nine are shared with the rest of the day and five (eggs, bacon, both sausages, hash brown) are breakfast-only, while eggs never co-occur with beef, lettuce or sour cream anywhere on the menu.

Thresholding the item graph (Figure 6) reveals a sharp transition. At Jaccard $\theta = 0.40$ the largest component still holds 53 items in 18 components; at $\theta = 0.45$ it holds 29; at 0.55, 14. Just below the transition the items with the highest betweenness are the Melty Pepper Jack Steak Burrito (0.221), which alone shares rice, nacho cheese and sour cream with the Loaded system and steak plus three-cheese with the quesadillas; the Breakfast Steak Quesadilla (0.148), which bridges breakfast to lunch; Crunchwrap Supreme (0.131) and Beefy 5-Layer Burrito (0.112), which carry both lettuce-cheddar and nacho-cheese-sour-cream and thus bridge Fresh to Loaded; and the Quesarito (0.095).

5.6 Linkage, motifs and alleles

Among pairs of components each in three or more items, 120 show positive linkage (lift > 1 with at least three co-occurrences). The strongest are the Cantina block: the two salsa packets together (lift 15.2, four items, $p = 3.9 \times 10^{-6}$), Avocado Verde Salsa with slow-roasted chicken (12.7, five items, $p = 3.3 \times 10^{-7}$), avocado ranch with purple cabbage (12.7), and pico with purple cabbage (9.5). This block behaves as an operon: introduced as a unit in 2024, it has never been split. Bacon and eggs (lift 6.3, $p = 4.3 \times 10^{-5}$) and gordita flatbread with spicy ranch (11.4) are the other strong linkages. Twenty-nine maximal filling motifs recur in four or more items; the most frequent are the Supreme motif (cheddar, lettuce, sour cream, beef, tomato; four items on four vehicles), the Cantina motif (avocado ranch, cheddar, lettuce, pico, cabbage; four items) and the Loaded motif (nacho cheese, sour cream, beef, three-cheese; five items).

More diagnostic than what co-occurs is what never does. Of the pairs in which both components appear in at least five items, 107 have zero co-occurrence; 28 of these have an

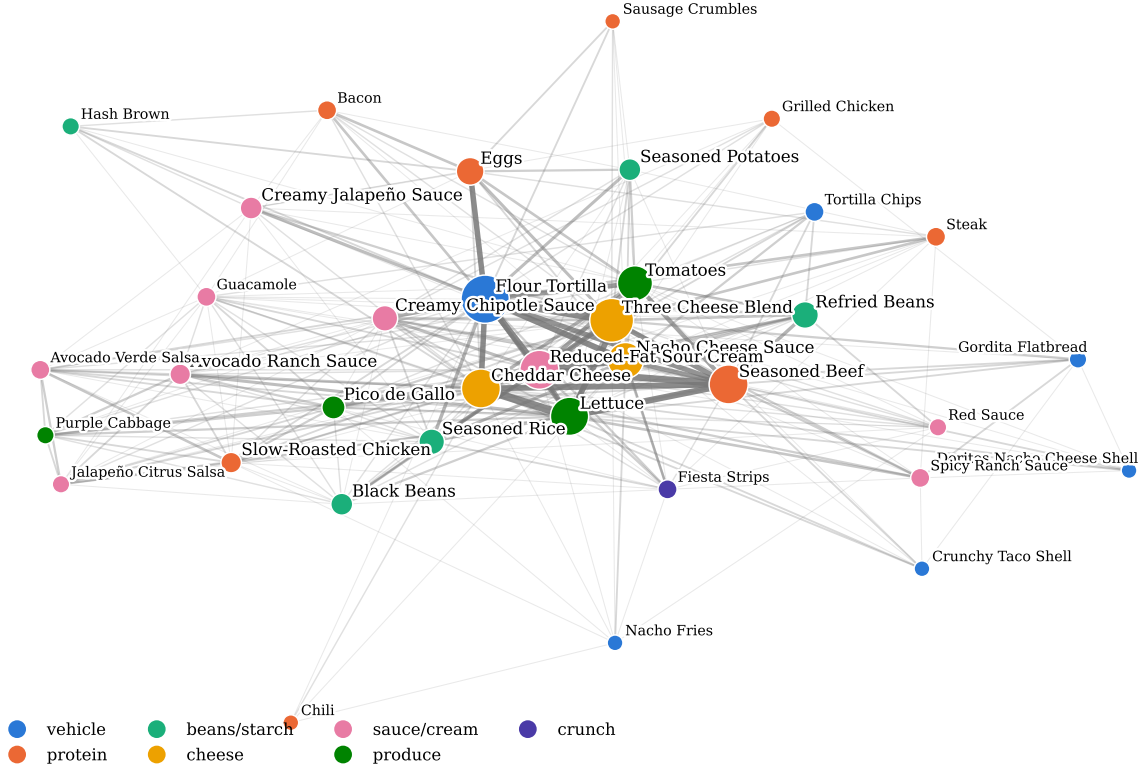


Figure 3: Component co-occurrence network (components in ≥ 3 items), spring layout. Node area is proportional to item count; edge width to the number of shared items; color is kitchen role.

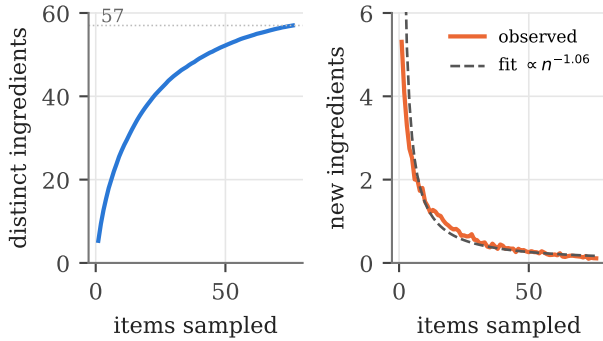


Figure 4: Pangenome rarefaction, mean of 300 random orderings. Left: distinct components after n items. Right: components new at step n , with the fitted power law $kn^{-\alpha}$, $\alpha = 1.06$.

expected co-occurrence of at least 1.5 under independence, and four are individually significant at $p < 0.05$ by Fisher’s test: eggs with sour cream ($E = 4.0$, $p = 0.007$), eggs with seasoned beef ($E = 4.0$, $p = 0.007$), eggs with lettuce ($E = 3.8$, $p = 0.014$) and flour tortilla with tortilla chips ($E = 2.6$, $p = 0.024$). With 76 genomes, most individual exclusions cannot reach significance; their structure is nonetheless clear when read as a group. Seasoned beef and black beans never co-occur ($E = 2.3$): black beans are the vegetarian allele of the protein slot, substituted in the Chalupa, Crunchwrap

and Grilled Cheese Burrito. Diced tomatoes and pico de gallo never co-occur ($E = 2.1$): two alleles of the tomato slot. Sour cream, creamy jalapeño sauce and spicy ranch are pairwise exclusive ($E = 2.3, 1.6, 0.5$): three alleles of a single “cool cream” slot. Eggs are separated from the lunch octet by daypart; tortilla and chips by physics. The kitchen enforces a grammar of one protein, one tomato form, one cool cream, and it does so without any published rule.

5.7 Predicted links and the candidate landscape

Link prediction on the projection scores 642 never-co-occurring pairs. After removing same-role pairs, the highest resource-allocation scores belong to rice with tomato (0.72; expected co-occurrence 2.6), cheddar with fiesta strips (0.70; 1.6), creamy jalapeño with tomato (0.66), purple cabbage with three-cheese blend (0.58), gordita flatbread with tomato (0.57), slow-roasted chicken with tomato (0.51), gordita flatbread with sour cream (0.51), slow-roasted chicken with refried beans (0.50), steak with cheddar (0.47) and steak with nacho cheese sauce (0.46). Almost every high-scoring pair joins a component characteristic of one flavor system to a component characteristic of another. The network is, in effect, predicting that the systems should mix more than they do.

Exhaustive enumeration under the grammar of Section 4.6 yields 1 205 088 candidate items. Their distribution over nov-

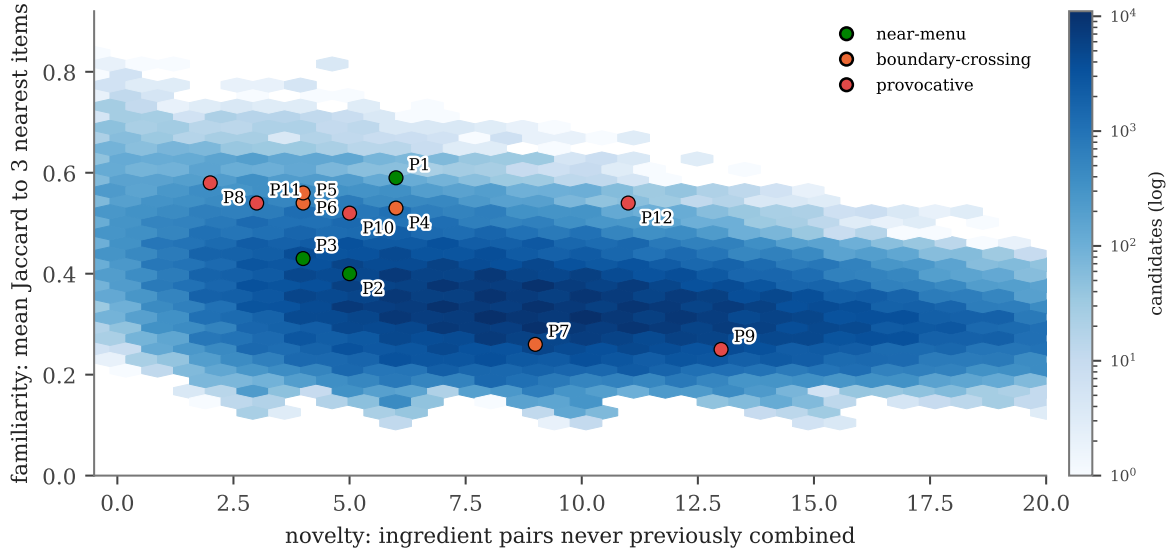


Figure 5: Familiarity–novelty landscape of 1 205 088 grammar-valid candidate items (hexbin, log count). Familiarity is the mean Jaccard similarity to the three nearest existing items; novelty is the number of component pairs never previously combined on the menu. P1–P12 are the proposals of Table 1.

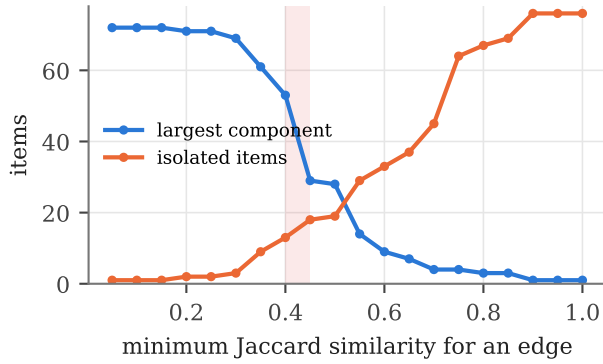


Figure 6: Percolation of the item similarity graph. The shaded band marks the transition between $\theta = 0.40$ (largest component 53) and $\theta = 0.45$ (29).

elty is unimodal with mode at eight never-seen pairs; 5012 candidates (0.4%) introduce no new pairwise combination at all, and a further 11 784 introduce exactly one. These are the menu’s adjacent possible: items every pair of whose components already coexists somewhere on the menu, yet which have never been assembled. Figure 5 plots familiarity against novelty for all candidates, with the twelve proposals of Section 6 overlaid. The upper envelope of the cloud is a Pareto front; the most familiar candidate at every novelty level up to eleven is a variant of the Grilled Cheese Burrito or Cheesy Double Beef Burrito, because large, heavily loaded items have many neighbors. Hand selection is therefore required to obtain proposals that are novel in *kind* and not merely in count, and the enumeration serves as a calibration surface rather than a generator.

6 Proposed products

Table 1 lists twelve proposals in three tiers, each scored by the same functions applied to the candidate landscape: nearest existing item and Jaccard similarity to it, familiarity, novelty, and the flavor systems whose signature components it draws on. The selection principles follow directly from the results. Near-menu proposals fill empty cells of the vehicle $\times$ protein grid identified in Section 5.1 or realize top-ranked predicted links without crossing a system boundary. Boundary-crossing proposals deliberately combine signature components of two systems, on the evidence that the menu’s few existing cross-system items (Crunchwrap Supreme, Beefy 5-Layer, Burrito Supreme) are also among its most durable. Provocative proposals break an allele rule, the daypart barrier, or the dessert–savory boundary, which is the only boundary in the corpus across which no gene has ever transferred.

Near-menu tier. **P1 Cantina Crunchwrap** fills the single most conspicuous gap in the vehicle $\times$ protein grid: slow-roasted chicken, the operator’s flagship 2024 protein, has never been placed on a tostada. Six of its component pairs are new, but all six involve the tostada, and its nearest neighbor is at Jaccard 0.73; it is a Cantina Chicken Burrito with a crunch layer. **P2 Steakhouse Gordita Crunch** closes the “no steak on any crunchy shell” gap and realizes two predicted links (steak with cheddar-family cheese, gordita with tomato) in one item. **P3 Fiesta Potatoes Supreme** realizes the predicted guacamole–potato link ($E = 0.46$, never observed) by applying the operator’s own Supreme and Grande operators to a side it already sells.

Boundary-crossing tier. **P4 Crunchwrap Unwrapped** is the top two predicted links (rice with tomato; cheddar with fi-

Table 1: Twelve proposed products with structural scores. Fam. is mean Jaccard to the three nearest existing items; Nov. is the number of component pairs never previously combined on the menu; Systems lists the flavor systems whose signature components are present.

#	Tier	Working name	Composition	Nearest item (<i>J</i>)	Fam.	Nov.	Systems
P1	Near	Cantina Crunchwrap	Flour tortilla, tostada shell, slow-roasted chicken, avocado ranch, creamy chipotle, lettuce, purple cabbage, pico, cheddar	Cantina Chicken Burrito (0.73)	0.60	6	Fresh, Loaded
P2	Near	Steakhouse Gordita Crunch	Gordita flatbread, crunchy shell, steak, three-cheese, pepper jack sauce, lettuce, tomato	Fresca Pepper Jack Steak Burrito (0.50)	0.40	5	Melt, Fresh
P3	Near	Fiesta Potatoes Supreme	Seasoned potatoes, nacho cheese, sour cream, guacamole, pico, fiesta strips	Cheesy Fiesta Potatoes (0.50)	0.43	4	Fresh, Loaded
P4	Cross	Crunchwrap Unwrapped	Crispy tortilla bowl, rice, beef, refried beans, lettuce, tomato, cheddar, sour cream, fiesta strips	Mini Taco Salad (0.60)	0.52	6	Fresh, Loaded
P5	Cross	Mexican Pizza Fries	Nacho Fries, Mexican Pizza sauce, beef, refried beans, three-cheese, tomato	Mexican Pizza (0.71)	0.56	4	Melt, Fresh, Loaded
P6	Cross	Melty Steak Quesarito	Flour tortilla, three-cheese, steak, pepper jack sauce, rice, nacho cheese, creamy chipotle	Melty Pepper Jack Steak Burrito (0.62)	0.53	4	Melt, Loaded
P7	Cross	Chili Crunch Tostada	Tostada shell, chili, cheddar, onions, Flamin' Hot Fritos, sour cream	Chili Cheese Burrito (0.29)	0.26	9	Fresh, Loaded
P8	Wild	The All-Dayer	Flour tortilla, eggs, seasoned beef, seasoned potatoes, nacho cheese, creamy chipotle	Beefy Potato Loaded Griller (0.71)	0.58	2	Melt, Loaded ^a
P9	Wild	Mango Verde Steak Taco	White-corn shell, steak, mango salsa, avocado verde salsa, purple cabbage, three-cheese, Tajín	Cantina Chicken Crispy Taco (0.30)	0.25	13	Melt, Fresh
P10	Wild	Naked Loaded Chalupa	Crispy chicken shell, rice, nacho cheese, three-cheese, fiesta strips, creamy chipotle	Grilled Cheese Burrito (0.56)	0.47	5	Melt, Loaded
P11	Wild	Cuatro Queso Quesarito	Flour tortilla, three-cheese, cheddar, nacho cheese, pepper jack sauce, beef, rice	Quesarito (0.56)	0.56	3	Melt, Fresh, Loaded
P12	Wild	The Cinnawrap	Flour tortilla, eggs, bacon, hash brown, cheddar, cinnamon twists, Cinnabon frosting	Breakfast Crunchwrap – Bacon (0.62)	0.54	11	Melt, Fresh, Sweet

^a Violates the daypart barrier (eggs with seasoned beef), the only proposal that breaks an observed exclusion rather than filling an empty cell.

esta strips) delivered as a bowl: the Supreme motif over rice, a Fresh item on a Loaded base. **P5 Mexican Pizza Fries** crosses all three systems and is the one proposal for which the operator has itself signalled appetite, having announced Mexican Pizza Empanadas for 2026; fries are a cheaper vehicle than

a new pastry. **P6 Melty Steak Quesarito** is built around the menu's most important bridge item and pushes it further into the Loaded system, adding nacho cheese and creamy chipotle to the pepper-jack steak base; it realizes the steak–nacho-cheese link (RA 0.46). **P7 Chili Crunch Tostada** takes the

highest-leverage low-frequency component, chili, and gives it a vehicle of its own for the first time; nine new pairs and a familiarity of 0.26 make it the most structurally novel item in the crossing tier, a Frito-pie tostada.

Provocative tier. **P8 The All-Dayer** breaks the strongest exclusion in the corpus. Eggs and seasoned beef have an expected co-occurrence of 4.0 and an observed count of zero; putting them in a Loaded Griller with potatoes and nacho cheese is structurally the smallest possible violation (two new pairs) with the largest expected surprise. **P9 Mango Verde Steak Taco** crosses the two most recent flavor introductions, the Cantina salsas and the Tajín–mango pairing, on the white-corn shell; thirteen new pairs and familiarity 0.25 place it at the outer edge of the landscape, and it is the proposal most likely to read as a genuinely new dish rather than a remix. **P10 Naked Loaded Chalupa** uses the crispy chicken shell, the menu’s only protein-as-vehicle, as the container for a Grilled Cheese Burrito filling; the enumeration itself surfaced this configuration on the Pareto front at seven new pairs. **P11 Cuatro Queso Quesarito** violates no allele rule but stacks all four cheese components in one item, something no existing item does with more than two; it is a maximal Melt item wearing a Loaded vehicle. **P12 The Cinnawrap** is the only proposal that crosses the dessert–savory boundary, the one boundary in the corpus with zero gene flow across it: a bacon-and-egg Breakfast Crunchwrap whose crunch layer is cinnamon twists and whose sauce is Cinnabon frosting. Eleven of its pairs are new, but its nearest neighbor is at 0.62, because everything except the two dessert genes is a menu item already. It is a sweet-and-salty breakfast sandwich, a category that exists elsewhere in quick service and nowhere on this menu.

7 Discussion

The menu as a closed pangenome. The most consequential result is the saturation of the ingredient vocabulary. A closed pangenome means the operator’s design space is, to a first approximation, fixed, and innovation is combinatorial. That has two implications. First, the search space is enumerable: we enumerated over a million grammar-valid items from the current pantry, and the grammar could be relaxed further. Second, the differentiating investment is in vehicles, not fillings. Every one of the six most distinctive current items owes its distance to a vehicle, and the 2026 pipeline continues the pattern. The operator’s own product language (Crunchwrap, Quesarito, Chalupa, Gordita, Stacker) is a language of containers.

Categories versus systems. The weak agreement between the operator’s categories and the recovered structure is not a criticism of the menu board, which is organized for ordering, but it matters for design. A new item’s likely reception is better predicted by the flavor system it belongs to than by the category it will be filed under. The Cantina line is the clearest example: it is marketed as a chicken line and is structurally a

Fresh-system line, and its extension into Loaded territory (P1, and the operator’s own Queso Cracked Cantina Bowl) is the natural next move under the systems view.

The Crunchwrap principle. The menu’s cross-system items are few and old. Crunchwrap Supreme (2005), Beefy 5-Layer (2009) and Burrito Supreme are the items that carry both Fresh and Loaded signatures, and they are among the operator’s most durable products. Link prediction independently ranks cross-system pairs highest. We conjecture, and cannot yet test, that cross-system items have higher survival on the menu than within-system items, and that the “surprising reception” the present exercise was asked to engineer is precisely the reception of a well-formed system crossing. Seven of the twelve proposals cross at least two systems for this reason.

Alleles as design constraints. The exclusion structure is the most practically useful output for anyone generating candidates. It says which substitutions are free (the protein slot, the tomato slot, the cool-cream slot) and therefore which permutations the operator will consider ordinary, and it says which pairs have never been tried and therefore where the surprise is. Of our proposals only P8 breaks an allele; the rest respect the grammar and find their novelty in vehicles and system crossings, which is where the corpus indicates the operator itself finds novelty.

8 Limitations

The corpus is one menu on one day, in one market; limited-time offers and regional items are excluded, and the operator’s history of items is not modelled, so “lineage” here is structural similarity rather than documented descent. Component presence is binary; quantities, preparation and layering order, which matter for taste, are not represented, and two items with identical component sets on different vehicles are treated as maximally similar in filling space. Ten items were decomposed from marketing descriptions rather than ingredient lists. The Heaps exponent is estimated from 76 genomes and is sensitive to the fitting convention. Louvain is stochastic and modularity is modest (0.21); the three-system partition was stable across the seeds we tried but alternative partitions with similar modularity exist. Fisher tests on a 76-item corpus have little power for individual pairs, and we have interpreted the exclusion structure as a group. Familiarity and novelty are structural proxies; neither is a measure of taste, and no sensory testing was performed. The product names are working names and have not been checked against trademark registers.

9 Conclusion

A quick-service menu of 76 items resolves into 57 components, of which eight fill more than half of all filling slots and 25 fill 82 percent; two thirds of items lie within two edits of another item; the ingredient vocabulary is saturated; and

the menu’s real organization is three flavor systems that its categories obscure. The kitchen enforces an implicit grammar of mutually exclusive alleles, and the network predicts that its systems should mix more than they do. The frontier is large—over a million grammar-valid items from the current pantry, five thousand of which introduce no new pairwise combination at all—and the twelve products proposed here are chosen from it to span the range from a Cantina Crunchwrap the operator could ship next quarter to a Cinnawrap that would, for the first time, move a gene across the dessert boundary.

Disclaimer. This manuscript is an independent academic study. The authors are not affiliated with, employed by, sponsored by, or otherwise connected to Taco Bell Corp., Yum! Brands, Inc., or any of their subsidiaries, franchisees or licensees, and no such entity has reviewed, approved or endorsed this work. TACO BELL, CRUNCHWRAP, CHALUPA, GORDITA, QUESARITO, CANTINA CHICKEN, LUXE, NACHO FRIES, LIVE MÁS and all related product names, menu descriptions and trade dress referenced herein are trademarks or registered trademarks of Taco Bell Corp. and/or Yum! Brands, Inc. DORITOS and FRITOS are trademarks of Frito-Lay North America, Inc.; CINNABON is a trademark of Cinnabon Franchisor SPV LLC; TAJÍN is a trademark of Industrias Tajín, S.A. de C.V. All such marks are the sole property of their respective owners and are used in this manuscript solely for the nominative purpose of identifying the publicly available products analysed; no license, sponsorship or endorsement is implied. Product concepts and working names proposed in Section 6 are hypothetical, have not been evaluated for trademark availability, and are offered for scholarly discussion only. All menu data were obtained from publicly accessible web pages as described in Section 3 and are reported as observed on the stated date; the authors make no representation as to their current accuracy.

Reproducibility. The component matrix, the analysis scripts (Python; NetworkX, SciPy, NumPy, Matplotlib) and the enumeration code are available from the corresponding author. Two interactive companion reports, *The Taco Bell Genome* and *The Taco Bell Phylogeny*, present the same analyses with hover-level detail.

References

- L. A. Adamic and E. Adar. Friends and neighbors on the Web. *Social Networks*, 25(3):211–230, 2003.
- Y.-Y. Ahn, S. E. Ahnert, J. P. Bagrow, and A.-L. Barabási. Flavor network and the principles of food pairing. *Scientific Reports*, 1:196, 2011.
- V. D. Blondel, J.-L. Guillaume, R. Lambiotte, and E. Lefebvre. Fast unfolding of communities in large networks. *Journal of Statistical Mechanics*, P10008, 2008.
- L. C. Freeman. A set of measures of centrality based on betweenness. *Sociometry*, 40(1):35–41, 1977.
- H. S. Heaps. *Information Retrieval: Computational and Theoretical Aspects*. Academic Press, 1978.
- L. Hubert and P. Arabie. Comparing partitions. *Journal of Classification*, 2:193–218, 1985.
- P. Jaccard. Étude comparative de la distribution florale dans une portion des Alpes et du Jura. *Bulletin de la Société Vaudoise des Sciences Naturelles*, 37:547–579, 1901.
- M. L. Kasavana and D. I. Smith. *Menu Engineering: A Practical Guide to Menu Analysis*. Hospitality Publications, 1982.
- D. Liben-Nowell and J. Kleinberg. The link-prediction problem for social networks. *Journal of the American Society for Information Science and Technology*, 58(7):1019–1031, 2007.
- D. Medini, C. Donati, H. Tettelin, V. Masignani, and R. Rappuoli. The microbial pan-genome. *Current Opinion in Genetics & Development*, 15(6):589–594, 2005.
- M. E. J. Newman. Modularity and community structure in networks. *Proceedings of the National Academy of Sciences*, 103(23):8577–8582, 2006.
- M. E. J. Newman. *Networks: An Introduction*. Oxford University Press, 2010.
- R. R. Sokal and C. D. Michener. A statistical method for evaluating systematic relationships. *University of Kansas Science Bulletin*, 38:1409–1438, 1958.
- C.-Y. Teng, Y.-R. Lin, and L. A. Adamic. Recipe recommendation using ingredient networks. In *Proceedings of the 4th ACM Web Science Conference*, pages 298–307, 2012.
- H. Tettelin et al. Genome analysis of multiple pathogenic isolates of *Streptococcus agalactiae*: implications for the microbial “pan-genome.” *Proceedings of the National Academy of Sciences*, 102(39):13950–13955, 2005.
- H. Tettelin, D. Riley, C. Cattuto, and D. Medini. Comparative genomics: the bacterial pan-genome. *Current Opinion in Microbiology*, 11(5):472–477, 2008.
- T. Zhou, L. Lü, and Y.-C. Zhang. Predicting missing links via local information. *European Physical Journal B*, 71:623–630, 2009.