



# Flavor network and the principles of food pairing

## SUBJECT AREAS:

STATISTICAL PHYSICS,  
THERMODYNAMICS AND  
NONLINEAR DYNAMICS  
APPLIED PHYSICS  
SYSTEMS BIOLOGY  
STATISTICS

Received  
18 October 2011

Accepted  
24 November 2011

Published  
15 December 2011

Correspondence and  
requests for materials  
should be addressed to  
S.E.A. (sea31@cam.  
ac.uk) or A.L.B. (alb@  
neu.edu)

\* These authors  
contributed equally to  
this work.

Yong-Yeol Ahn<sup>1,2,3\*</sup>, Sebastian E. Ahnert<sup>1,4\*</sup>, James P. Bagrow<sup>1,2</sup> & Albert-László Barabási<sup>1,2</sup>

<sup>1</sup>Center for Complex Network Research, Department of Physics Northeastern University, Boston, MA 02115, <sup>2</sup>Center for Cancer Systems Biology Dana-Farber Cancer Institute, Harvard University, Boston, MA 02115, <sup>3</sup>School of Informatics and Computing Indiana University, Bloomington, IN 47408, <sup>4</sup>Theory of Condensed Matter, Cavendish Laboratory, University of Cambridge, Cambridge CB3 0HE, UK.

The cultural diversity of culinary practice, as illustrated by the variety of regional cuisines, raises the question of whether there are any general patterns that determine the ingredient combinations used in food today or principles that transcend individual tastes and recipes. We introduce a flavor network that captures the flavor compounds shared by culinary ingredients. Western cuisines show a tendency to use ingredient pairs that share many flavor compounds, supporting the so-called food pairing hypothesis. By contrast, East Asian cuisines tend to avoid compound sharing ingredients. Given the increasing availability of information on food preparation, our data-driven investigation opens new avenues towards a systematic understanding of culinary practice.

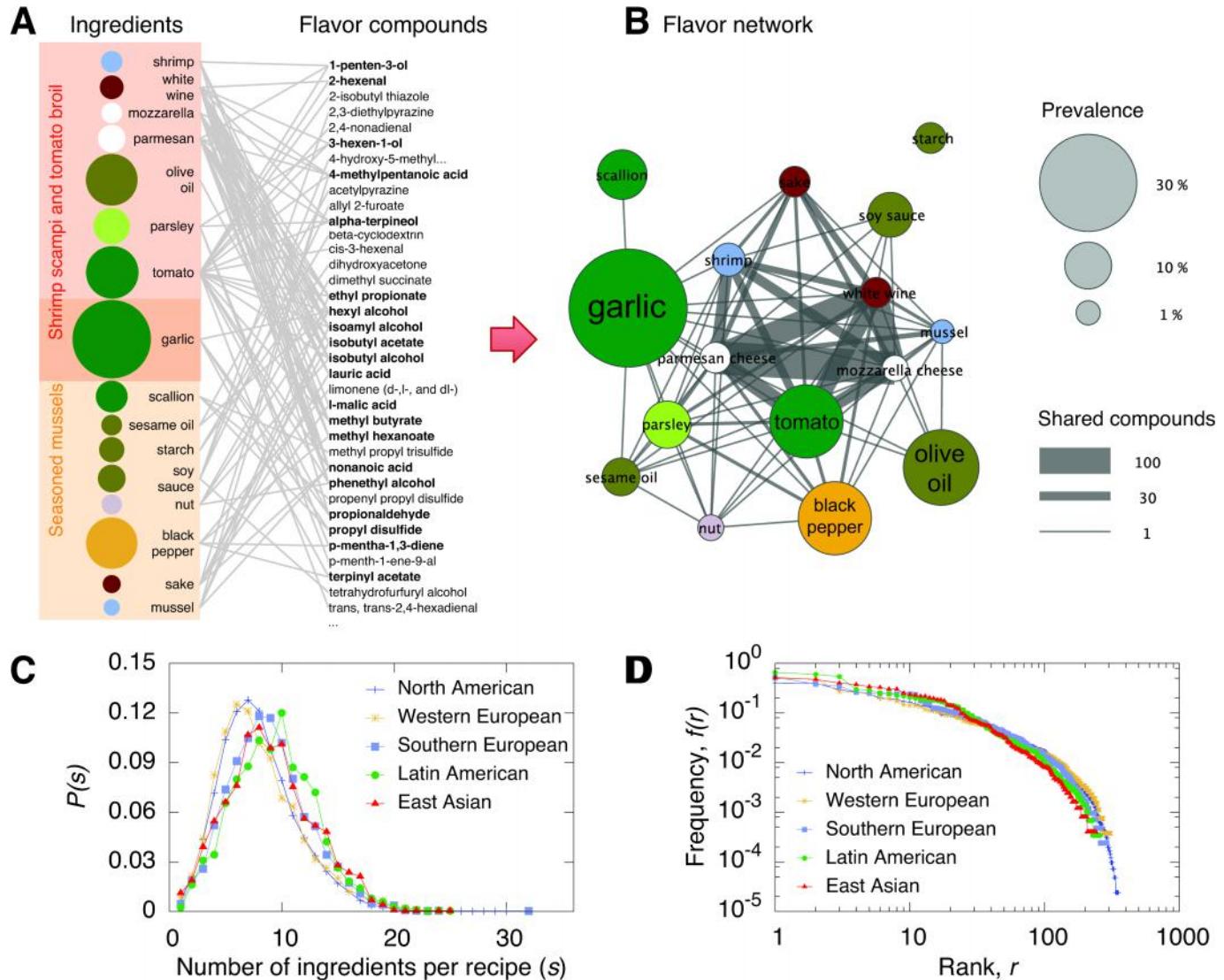
As omnivores, humans have historically faced the difficult task of identifying and gathering food that satisfies nutritional needs while avoiding foodborne illnesses<sup>1</sup>. This process has contributed to the current diet of humans, which is influenced by factors ranging from an evolved preference for sugar and fat to palatability, nutritional value, culture, ease of production, and climate<sup>1–9</sup>. The relatively small number of recipes in use ( $\sim 10^6$ , e.g. <http://cookpad.com>) compared to the enormous number of potential recipes ( $> 10^{15}$ , see Supplementary Information Sec S1.2), together with the frequent recurrence of particular combinations in various regional cuisines, indicates that we are exploiting but a tiny fraction of the potential combinations. Although this pattern itself can be explained by a simple evolutionary model<sup>10</sup> or data-driven approaches<sup>11</sup>, a fundamental question still remains: are there any quantifiable and reproducible principles behind our choice of certain ingredient combinations and avoidance of others?

Although many factors such as colors, texture, temperature, and sound play an important role in food sensation<sup>12–15</sup>, palatability is largely determined by flavor, representing a group of sensations including odors (due to molecules that can bind olfactory receptors), tastes (due to molecules that stimulate taste buds), and freshness or pungency (trigeminal senses)<sup>16</sup>. Therefore, the flavor compound (chemical) profile of the culinary ingredients is a natural starting point for a systematic search for principles that might underlie our choice of acceptable ingredient combinations.

A hypothesis, which over the past decade has received attention among some chefs and food scientists, states that ingredients sharing flavor compounds are more likely to taste well together than ingredients that do not<sup>17</sup> (also see <http://www.foodpairing.com>). This food pairing hypothesis has been used to search for novel ingredient combinations and has prompted, for example, some contemporary restaurants to combine white chocolate and caviar, as they share trimethylamine and other flavor compounds, or chocolate and blue cheese that share at least 73 flavor compounds. As we search for evidence supporting (or refuting) any ‘rules’ that may underlie our recipes, we must bear in mind that the scientific analysis of any art, including the art of cooking, is unlikely to be capable of explaining every aspect of the artistic creativity involved. Furthermore, there are many ingredients whose main role in a recipe may not be only flavoring but something else as well (e.g. eggs’ role to ensure mechanical stability or paprika’s role to add vivid colors). Finally, the flavor of a dish owes as much to the mode of preparation as to the choice of particular ingredients<sup>12,18,19</sup>. However, our hypothesis is that, given the large number of recipes we use in our analysis (56,498), such factors can be systematically filtered out, allowing for the discovery of patterns that may transcend specific dishes or ingredients.

Here we introduce a network-based approach to explore the impact of flavor compounds on ingredient combinations. Efforts by food chemists to identify the flavor compounds contained in most culinary ingredients allows us to link each ingredient to 51 flavor compounds on average<sup>20</sup><sup>1</sup>. We build a bipartite network<sup>21–26</sup>

<sup>1</sup>While finalizing this manuscript, an updated edition (6th Ed.) of Fenaroli’s handbook of flavor ingredients has been released.



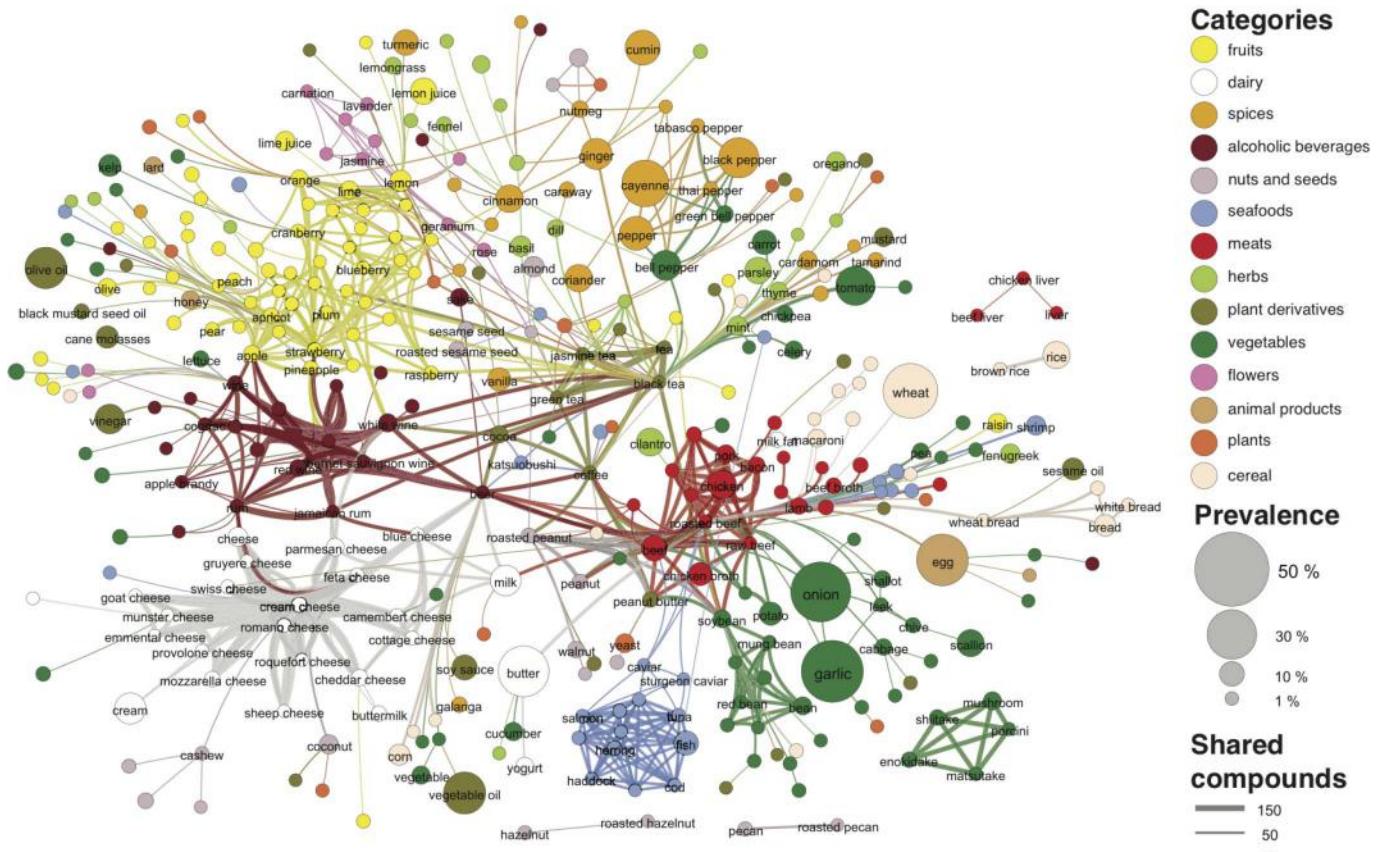
**Figure 1 | Flavor network.** (A) The ingredients contained in two recipes (left column), together with the flavor compounds that are known to be present in the ingredients (right column). Each flavor compound is linked to the ingredients that contain it, forming a bipartite network. Some compounds (shown in boldface) are shared by multiple ingredients. (B) If we project the ingredient-compound bipartite network into the ingredient space, we obtain the *flavor network*, whose nodes are ingredients, linked if they share at least one flavor compound. The thickness of links represents the number of flavor compounds two ingredients share and the size of each circle corresponds to the prevalence of the ingredients in recipes. (C) The distribution of recipe size, capturing the number of ingredients per recipe, across the five cuisines explored in our study. (D) The frequency-rank plot of ingredients across the five cuisines show an approximately invariant distribution across cuisines.

consisting of two different types of nodes: (i) 381 ingredients used in recipes throughout the world, and (ii) 1,021 flavor compounds that are known to contribute to the flavor of each of these ingredients (Fig. 1A). A projection of this bipartite network is the *flavor network* in which two nodes (ingredients) are connected if they share at least one flavor compound (Fig. 1B). The weight of each link represents the number of shared flavor compounds, turning the flavor network into a weighted network<sup>27,22,23</sup>. While the compound concentration in each ingredient and the detection threshold of each compound should ideally be taken into account, the lack of systematic data prevents us from exploring their impact (see Sec S1.2 on data limitations).

Since several flavor compounds are shared by a large number of ingredients, the resulting flavor network is too dense for direct visualization (average degree  $\langle k \rangle \approx 214$ ). We therefore use a backbone extraction method<sup>28,29</sup> to identify the statistically significant links for each ingredient given the sum of weights characterizing the particular node (Fig. 2), see SI for details). Not surprisingly, each module in the network corresponds to a distinct food class such as meats (red)

or fruits (yellow). The links between modules inform us of the flavor compounds that hold different classes of foods together. For instance, fruits and dairy products are close to alcoholic drinks, and mushrooms appear isolated, as they share a statistically significant number of flavor compounds only with other mushrooms.

The flavor network allows us to reformulate the food pairing hypothesis as a topological property: do we more frequently use ingredient pairs that are strongly linked in the flavor network or do we avoid them? To test this hypothesis we need data on ingredient combinations preferred by humans, information readily available in the current body of recipes. For generality, we used 56,498 recipes provided by two American repositories ([epicurious.com](http://epicurious.com) and [allrecipes.com](http://allrecipes.com)) and to avoid a distinctly Western interpretation of the world's cuisine, we also used a Korean repository ([menupan.com](http://menupan.com)). The recipes are grouped into geographically distinct cuisines (North American, Western European, Southern European, Latin American, and East Asian; see Fig. 1 and Table S2). The average number of ingredients used in a recipe is around eight, and the overall distribution is bounded (Fig. 1C), indicating that recipes with a very large



**Figure 2 | The backbone of the flavor network.** Each node denotes an ingredient, the node color indicates food category, and node size reflects the ingredient prevalence in recipes. Two ingredients are connected if they share a significant number of flavor compounds, link thickness representing the number of shared compounds between the two ingredients. Adjacent links are bundled to reduce the clutter. Note that the map shows only the statistically significant links, as identified by the algorithm of Refs.<sup>28,29</sup> for  $p$ -value 0.04. A drawing of the full network is too dense to be informative. We use, however, the full network in our subsequent measurements.

or very small number of ingredients are rare. By contrast, the popularity of specific ingredients varies over four orders of magnitude, documenting huge differences in how frequently various ingredients are used in recipes (Fig. 1D), as observed in<sup>10</sup>. For example, jasmine tea, Jamaican rum, and 14 other ingredients are each found in only a single recipe (see SI S1.2), but egg appears in as many as 20,951, more than one third of all recipes.

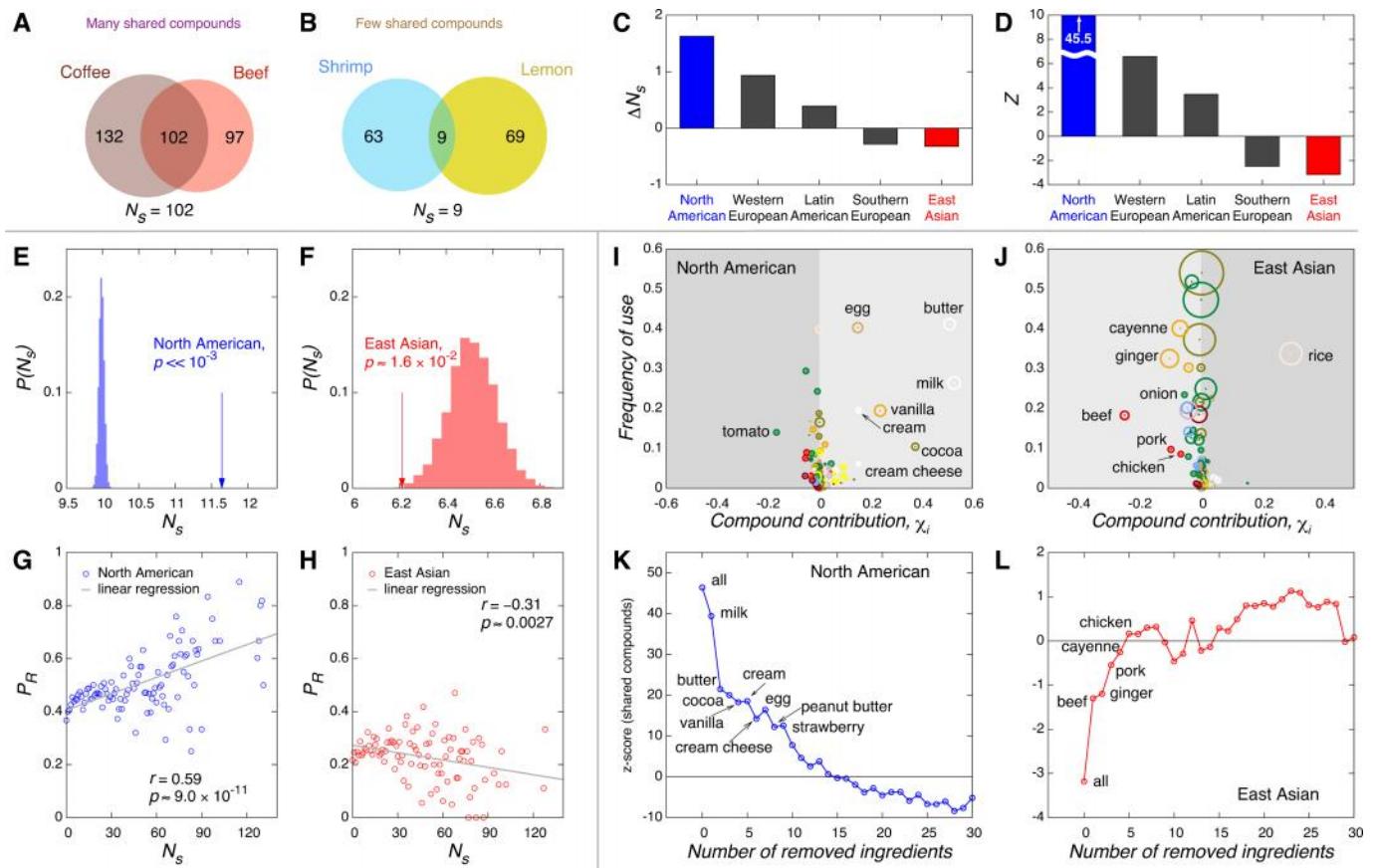
## Results

Figure 3D indicates that North American and Western European cuisines exhibit a statistically significant tendency towards recipes whose ingredients share flavor compounds. By contrast, East Asian and Southern European cuisines avoid recipes whose ingredients share flavor compounds (see Fig. 3D for the Z-score, capturing the statistical significance of  $\Delta N_s$ ). The systematic difference between the East Asian and the North American recipes is particularly clear if we inspect the  $P(N_s^{\text{rand}})$  distribution of the randomized recipe dataset, compared to the observed number of shared compounds characterizing the two cuisines,  $N_s$ . This distribution reveals that North American dishes use far more compound-sharing pairs than expected by chance (Fig. 3E), and the East Asian dishes far fewer (Fig. 3F). Finally, we generalize the food pairing hypothesis by exploring if ingredient pairs sharing more compounds are more likely to be used in specific cuisines. The results largely correlate with our earlier observations: in North American recipes, the more compounds are shared by two ingredients, the more likely they appear in recipes. By contrast, in East Asian cuisine the more flavor compounds two

ingredients share, the less likely they are used together (Fig. 3G and 3H; see SI for details and results on other cuisines).

What is the mechanism responsible for these differences? That is, does Fig. 3C through H imply that all recipes aim to pair ingredients together that share (North America) or do not share (East Asia) flavor compounds, or could we identify some compounds responsible for the bulk of the observed effect? We therefore measured the contribution  $\chi_i$  (see Methods) of each ingredient to the shared compound effect in a given cuisine  $c$ , quantifying to what degree its presence affects the magnitude of  $\Delta N_s$ .

In Fig. 3I,J we show as a scatter plot  $\chi_i$  (horizontal axis) and the frequency  $f_i$  for each ingredient in North American and East Asian cuisines. The vast majority of the ingredients lie on the  $\chi_i = 0$  axis, indicating that their contribution to  $\Delta N_s$  is negligible. Yet, we observe a few frequently used outliers, which tend to be in the positive  $\chi_i$  region for North American cuisine, and lie predominantly in the negative region for East Asian cuisine. This suggests that the food pairing effect is due to a few outliers that are frequently used in a particular cuisine, e.g. milk, butter, cocoa, vanilla, cream, and egg in the North America, and beef, ginger, pork, cayenne, chicken, and onion in East Asia. Support for the definitive role of these ingredients is provided in Fig. 3K,L where we removed the ingredients in order of their positive (or negative) contributions to  $\Delta N_s$  in the North American (or East Asian) cuisine, finding that the z-score, which measures the significance of the shared compound hypothesis, drops below two after the removal of only 13 (5) ingredients from North American (or East Asian) cuisine (see SI S2.2.2). Note, however, that



**Figure 3 | Testing the food pairing hypothesis.** Schematic illustration of two ingredient pairs, the first sharing many more (A) and the second much fewer (B) compounds than expected if the flavor compounds were distributed randomly. (C,D) To test the validity of the food pairing hypothesis, we construct 10,000 random recipes and calculate  $\Delta N_s$ . We find that ingredient pairs in North American cuisines tend to share more compounds while East Asian cuisines tend to share fewer compounds than expected in a random recipe dataset. (E,F) The distributions  $P(N_s)$  for 10,000 randomized recipe datasets compared with the real values for East Asian and North American cuisine. Both cuisines exhibit significant  $p$ -values, as estimated using a  $z$ -test. (G,H) We enumerate every possible ingredient pair in each cuisine and show the fraction of pairs in recipes as a function of the number of shared compounds. To reduce noise, we only used data points calculated from more than 5 pairs. The  $p$ -values are calculated using a  $t$ -test. North American cuisine is biased towards pairs with more shared compounds while East Asian shows the opposite trend (see SI for details and results for other cuisines). Note that we used the full network, not the backbone shown in Fig. 2 to obtain these results. (I,J) The contribution and frequency of use for each ingredient in North American and East Asian cuisine. The size of the circles represents the relative prevalence  $p_i^c$ . North American and East Asian cuisine shows the opposite trends. (K,L) If we remove the highly contributing ingredients sequentially (from the largest contribution in North American cuisine and from the smallest contribution in East Asian cuisine), the shared compounds effect quickly vanishes when we removed five (East Asian) to fifteen (North American) ingredients.

these ingredients play a disproportionate role in the cuisine under consideration—for example, the 13 key ingredients contributing to the shared compound effect in North American cuisine appear in 74.4% of all recipes.

According to an empirical view known as “the flavor principle”<sup>30</sup>, the differences between regional cuisines can be reduced to a few key ingredients with specific flavors: adding soy sauce to a dish almost automatically gives it an oriental taste because Asians use soy sauce widely in their food and other ethnic groups do not; by contrast paprika, onion, and lard is a signature of Hungarian cuisine. Can we systematically identify the ingredient combinations responsible for the taste palette of a regional cuisine? To answer this question, we measure the *authenticity* of each ingredient ( $p_i^c$ ), ingredient pair ( $p_{ij}^c$ ), and ingredient triplet ( $p_{ijk}^c$ ) (see Methods). In Fig. 4 we organize the six most authentic single ingredients, ingredient pairs and triplets for North American and East Asian cuisines in a flavor pyramid. The rather different ingredient classes (as reflected by their color) in the two pyramids capture the differences between the two cuisines: North American food heavily relies on dairy products, eggs and wheat; by contrast, East Asian cuisine is dominated by plant derivatives like soy sauce, sesame oil, and rice and ginger. Finally, the two pyramids also illustrate the different affinities of the two regional

cuisines towards food pairs with shared compounds. The most authentic ingredient pairs and triplets in the North American cuisine share multiple flavor compounds, indicated by black links, but such compound-sharing links are rare among the most authentic combinations in East Asian cuisine.

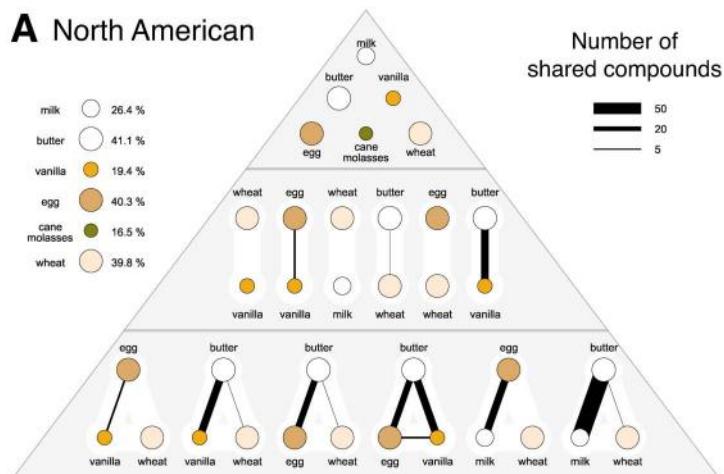
The reliance of regional cuisines on a few authentic ingredient combinations allows us to explore the ingredient-based relationship (similarity or dissimilarity) between various regional cuisines. For this we selected the six most authentic ingredients and ingredient pairs in each regional cuisine (i.e. those shown in Fig. 4A,B), generating a diagram that illustrates the ingredients shared by various cuisines, as well as singling out those that are unique to a particular region (Fig. 4C). We once again find a close relationship between North American and Western European cuisines and observe that when it comes to its signature ingredient combinations Southern European cuisine is much closer to Latin American than Western European cuisine (Fig. 4C).

## Discussion

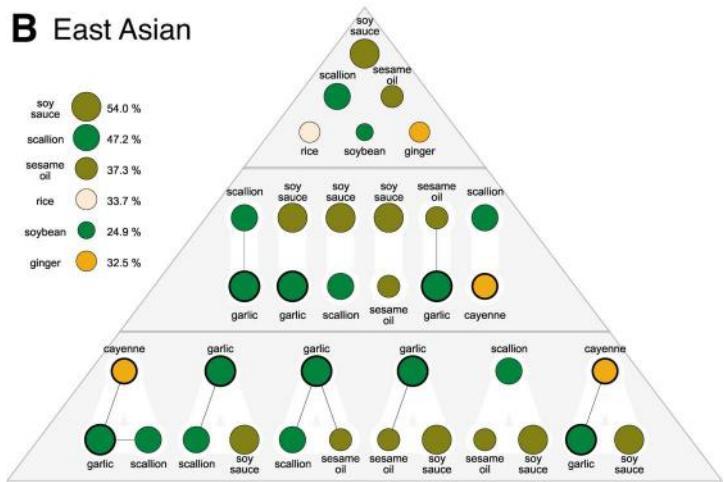
Our work highlights the limitations of the recipe data sets currently available, and more generally of the systematic analysis of food



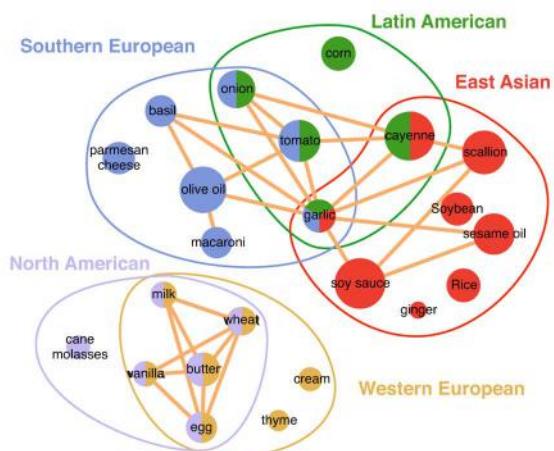
### A North American



### B East Asian



### C Co-occurrence in recipes



**Figure 4 | Flavor principles.** (A,B) Flavor pyramids for North American and East Asian cuisines. Each flavor pyramid shows the six most **authentic** ingredients (i.e. those with the largest  $p_i^c$ ), ingredient pairs (largest  $p_{ij}^c$ ), and ingredient triplets (largest  $p_{ijk}^c$ ). The size of the nodes reflects the abundance  $P_i$  of the ingredient in the recipes of the particular cuisine. Each color represents the category of the ingredient (see Fig. 2 for the color) and link thickness indicates the number of shared compounds. (C) The six most authentic ingredients and ingredient pairs used in specific regional cuisine. Node color represents cuisine and the link weight reflects the relative prevalence  $p_i^c$  of the ingredient pair.

preparation data. By comparing two editions of the same dataset with significantly different coverage, we can show that our results are robust against data incompleteness (see SI S1.1.2). Yet, better compound databases, mitigating the incompleteness and the potential

biases of the current data, could significantly improve our understanding of food. There is inherent ambiguity in the definition of a particular regional or ethnic cuisine. However, as discussed in SI S1.2, the correlation between different datasets, representing two distinct



perspectives on food (American and Korean), indicates that humans with different ethnic background have a rather consistent view on the composition of various regional cuisines.

Recent work by Kinouchi et al.<sup>10</sup> observed that the frequency-rank plots of ingredients are invariant across four different cuisines, exhibiting a shape that can be well described by a Zipf-Mandelbrot curve. Based on this observation, they model the evolution of recipes by assuming a copy-mutate process, leading to a very similar frequency-rank curve. The copy-mutate model provides an explanation for how an ingredient becomes a staple ingredient of a cuisine: namely, having a high value or being a founder ingredient. The model assigns each ingredient a random fitness value, which represents the ingredient's nutritional value, availability, flavor, etc. For example, it has been suggested that some ingredients are selected because of their antimicrobial properties<sup>6,7</sup>. The mutation phase of the model replaces less fit ingredients with fitter ones. Meanwhile, the copy mechanism keeps copying the *founder ingredients*—ingredients in the early recipes—and makes them abundant in the recipes regardless of their fitness value.

It is worthwhile to discuss the similarities and differences between the quantities we measured and the concepts of fitness and founders. First of all, prevalence ( $P_i^c$ ) and authenticity ( $p_i^c$ ) are empirically measured values while fitness is an intrinsic hidden variable. Among the list of highly prevalent ingredients we indeed find *old* ingredients—founders—that have been used in the same geographic region for thousands of years. At the same time, there are relatively new ingredients such as tomatoes, potatoes, and peppers that were introduced to Europe and Asia just a few hundred years ago. These new, but prevalent ingredients can be considered to have high fitness values. If an ingredient has a high level of authenticity, then it is prevalent in a cuisine while not so prevalent in all other cuisines.

Indeed, each culture has developed their own authentic ingredients. It may indicate that fitness can vary greatly across cuisines or that the stochasticity of recipe evolution diverge the recipes in different regions into completely different sets. More historical investigation will help us to estimate the fitness of ingredients and assess why we use the particular ingredients we currently do. The higher order fitness value suggested in Kinouchi et al. is very close to our concept of food pairing affinity.

Another difference in our results is the number of ingredients in recipes. Kinouchi et al. reported that the average number of ingredients per recipe varies across different cookbooks. While we also observed variation in the number of ingredients per recipe, the patterns we found were not consistent with those found by Kinouchi et al. For instance, the French cookbook has more ingredients per recipe than a Brazilian one, but in our dataset we find the opposite result. We believe that a cookbook cannot represent a whole cuisine, and that cookbooks with more sophisticated recipes will tend to have more ingredients per recipe than cookbooks with everyday recipes. As more complete datasets become available, sharper conclusions can be drawn regarding the size variation between cuisines.

Our contribution in this context is a study of the role that flavour compounds play in determining these fitness values. One possible interpretation of our results is that shared flavor compounds represent one of several contributions to fitness value, and that, while shared compounds clearly play a significant role in some cuisines, other contributions may play a more dominant role in other cuisines. The fact that recipes rely on ingredients not only for flavor but also to provide the final textures and overall structure of a given dish provides support for the idea that fitness values depend on a multitude of ingredient characteristics besides their flavor profile.

In summary, our network-based investigation identifies a series of statistically significant patterns that characterize the way humans choose the ingredients they combine in their food. These patterns manifest themselves to varying degree in different geographic regions: while North American and Western European dishes tend to combine

ingredients that share flavor compounds, East Asian cuisine avoids them. More generally this work provides an example of how the data-driven network analysis methods that have transformed biology and the social sciences in recent years can yield new insights in other areas, such as food science.

## Methods

**Shared compounds.** To test the hypothesis that the choice of ingredients is driven by an appreciation for ingredient pairs that share flavor compounds (i.e. those linked in Fig. 2), we measured the mean number of shared compounds in each recipe,  $N_s$ , comparing it with  $N_s^{\text{rand}}$  obtained for a randomly constructed reference recipe dataset. For a recipe  $R$  that contains  $n_R$  different ingredients, where each ingredient  $i$  has a set of flavor compounds  $C_i$ , the mean number of shared compounds

$$N_s(R) = \frac{2}{n_R(n_R - 1)} \sum_{i,j \in R, i \neq j} |C_i \cap C_j| \quad (1)$$

is zero if none of the ingredient pairs  $(i, j)$  in the recipe share any flavor compounds. For example, the 'mustard cream pan sauce' recipe contains chicken broth, mustard, and cream, none of which share any flavor compounds ( $N_s(R) = 0$ ) in our dataset. Yet,  $N_s(R)$  can reach as high as 60 for 'sweet and simple pork chops', a recipe containing apple, pork, and cheddar cheese (See Fig. 3A). To check whether recipes with high  $N_s(R)$  are statistically preferred (implying the validity of the shared compound hypothesis) in a cuisine  $c$  with  $N_c$  recipes, we calculate  $\Delta N_s = N_s^{\text{real}} - N_s^{\text{rand}}$ , where 'real' and 'rand' indicates real recipes and randomly constructed recipes respectively and  $N_s = \sum_R N_s(R)/N_c$  (see SI for details of the randomization process). This random reference (null model) controls for the frequency of a particular ingredient in a given regional cuisine, hence our results are not affected by historical, geographical, and climate factors that determine ingredient availability (see SI S1.1.2).

**Contribution.** The contribution  $\chi_i$  of each ingredient to the shared compound effect in a given cuisine  $c$ , quantifying to what degree its presence affects the magnitude of  $\Delta N_s$ , is defined by

$$\chi_i = \left( \frac{1}{N_c} \sum_{R \in c} \frac{2}{n_R(n_R - 1)} \sum_{j \neq i, j \in R} |C_i \cap C_j| \right) - \left( \frac{2f_i}{N_c \langle n_R \rangle} \frac{\sum_{j \in c} f_j |C_i \cap C_j|}{\sum_{j \in c} f_j} \right), \quad (2)$$

where  $f_i$  represents the ingredient  $i$ 's number of occurrence. An ingredient's contribution is positive (negative) if it increases (decreases)  $\Delta N_s$ .

**Authenticity.** we define the prevalence  $P_i^c$  of each ingredient  $i$  in a cuisine  $c$  as  $P_i^c = n_i^c / N_c$ , where  $n_i^c$  is the number of recipes that contain the particular ingredient  $i$  in the cuisine and  $N_c$  is the total number of recipes in the cuisine. The relative prevalence  $p_i^c = P_i^c - \langle P_i^c \rangle_{c \neq c}$  measures the authenticity—the difference between the prevalence of  $i$  in cuisine  $c$  and the average prevalence of  $i$  in all other cuisines. We can also identify ingredient pairs or triplets that are overrepresented in a particular cuisine relative to other cuisines by defining the relative pair prevalences  $P_{ij}^c = P_{ij}^c - \langle P_{ij}^c \rangle_{c \neq c}$  and triplet prevalences  $P_{ijk}^c = P_{ijk}^c - \langle P_{ijk}^c \rangle_{c \neq c}$ , with  $P_{ij}^c = n_{ij}^c / N_c$  and  $P_{ijk}^c = n_{ijk}^c / N_c$ .

1. Rozin, P. The selection of foods by rats, humans, and other animals. *Advances in the Study of Behavior* **7**, 21–76 (1976).
2. Pfaffman, C. *Olfaction and taste V*, chap. Phylogenetic origins of sweet sensitivity (Academic Press, New York, 1975).
3. Garcia, J. & Hanks, W. G. *Olfaction and taste V*, chap. The evolution of bitter and the acquisition of toxiphobia (Academic Press, New York, 1975).
4. Drewnowski, A. & Greenwood, M. R. C. Cream and sugar: human preferences for highfat foods. *Physiology & Behavior* **30**, 629–633 (1983).
5. Diamond, J. M. *Guns, germs, and steel: The fates of human societies* (W.W. Norton, New York, 1997).
6. Billing, J. & Sherman, P. W. Antimicrobial functions of spices: why some like it hot. *The Quarterly Review of Biology* **73**, 3–49 (1998).
7. Sherman, P. W. & Hash, G. A. Why vegetable recipes are not very spicy. *Evolution and Human Behavior* **22**, 147–163 (2001).
8. Harris, M. *Good to eat: riddles of food and culture* (Waveland Press, 1998).
9. Counihan, C. & van Esterik, P. (eds.) *Food and culture* (Routledge, 2007).
10. Kinouchi, O., Diez-Garcia, R. W., Holanda, A. J., Zambianchi, P. & Roque, A. C. The non-equilibrium nature of culinary evolution. *New Journal of Physics* **10**, 073020 (2008).
11. Teng, C.-Y., Lin, Y.-R. & Adamic, L. A. Recipe recommendation using ingredient networks (2011). ArXiv:1111.3919 [cs.SI].
12. This, H. *Molecular gastronomy: exploring the science of flavor* (Columbia University Press, 2005).
13. Johnson, J. & Clydesdale, F. Perceived sweetness and redness in colored sucrose solutions. *Journal of Food Science* **47**, 747–752 (1982).
14. Shankaer, M. U. & Levitan, C. A. Grape expectations: the role of cognitive influences in color-flavor interactions. *Conscious Cogn.* **19**, 380–390 (2010).



15. Zampini, M. & Spence, C. The role of auditory cues in modulating the perceived crispness and staleness of potato chips. *Journal of Sensory Studies* **19**, 347–363 (2005).
16. Breslin, P. A. S. & Beauchamp, G. K. Suppression of bitterness by sodium: variation among bitter taste stimuli. *Chem. Senses* **20**, 609–623 (1995).
17. Blumenthal, H. *The big fat duck cookbook* (Bloomsbury, London, 2008).
18. This, H. Molecular gastronomy, a scientific look at cooking. *Accounts of Chemical Research* **42**, 575–583 (2009).
19. McGee, H. *On food and cooking: the science and lore of the kitchen* (Scribner, 2004).
20. Burdock, G. A. *Fenaroli's handbook of flavor ingredients* (CRC Press, 2004), 5th edn.
21. Newman, M. E. J., Barabási, A.-L. & Watts, D. J. *The structure and dynamics of networks* (Princeton University Press, 2006).
22. Caldarelli, G. *Scale-free networks: complex webs in nature and technology* (Oxford University Press, USA, 2007).
23. Dorogovtsev, S. N. & Mendes, J. F. F. *Evolution of networks: from biological nets to the internet and WWW* (Oxford University Press, USA, 2003).
24. Albert, R. & Barabási, A.-L. Statistical mechanics of complex networks. *Rev. Mod. Phys.* **74**, 47 (2002).
25. Newman, M. E. J. The structure and function of complex networks. *SIAM Review* **45**, 167–256 (2003).
26. Dorogovtsev, S. N., Goltsev, A. V. & Mendes, J. F. F. Critical phenomena in complex networks. *Reviews of Modern Physics* **80**, 1275–61 (2008).
27. Barrat, A., Barthélémy, M., Pastor-Satorras, R. & Vespignani, A. The architecture of complex weighted networks. *Proc. Nat. Acad. Sci.* **111**, 3747 (2004).
28. Serrano, M. A., Boguñá, M. & Vespignani, A. Extracting the multiscale backbone of complex weighted networks. *Proceedings of the National Academy of Sciences* **106**, 6483–6488 (2009).
29. Lee, S. H., Kim, P.-J., Ahn, Y.-Y. & Jeong, H. Googling social interactions: web search engine based social network construction. *PLoS One* **5**, e11233 (2010).
30. Rozin, E. *The flavor-principle cookbook* (Hawthorn Books, Book Club Ed. edition, 1973).

## Acknowledgements

We thank M. I. Meirelles, S. Lehmann, D. Kudayarova, T. W. Lim, J. Baranyi, H. This for discussions. This work was supported by the James S. McDonnell Foundation 21st Century Initiative in Studying Complex Systems.

## Author contributions

YYA, SEA, and ALB designed research and YYA, SEA, and JPB performed research. All authors wrote and reviewed the manuscript.

## Additional information

Supplementary information accompanies this paper at <http://www.nature.com/scientificreports/>

**Competing financial interests:** The authors declare no competing financial interests.

**License:** This work is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-sa/3.0/>

**How to cite this article:** Ahn, Y.-Y., Ahnert, S.E., Bagrow, J.P. & Barabási, A.-L. Flavor network and the principles of food pairing. *Sci. Rep.* **1**, 196; DOI:10.1038/srep00196 (2011).