Local and regional food production diversity are positively associated with household dietary diversity in rural Africa

Thanh-Tung Nguyen [1] ✉ **& Matin Qaim** [1,2]

Undernutrition and low dietary quality remain widespread issues in Africa. As most rural households in the region are involved in farming, the diversification of own farm production could improve their access to nutritious foods. Here we use representative panel data from six African countries to estimate this effect across different scales. We show that farm production diversity is positively associated with household dietary diversity—yet the average magnitude of the association is small, depends on the specific measure of production diversity and increases with distance from urban centres. In all countries, markets and market access are more important for dietary diversity than own production. Because village-, town- and district-level production diversity are often positively associated with dietary diversity, higher diversity on each individual farm may not be required. The appropriate spatial scale should be considered when designing diversification strategies.

More than two billion people worldwide suffer from undernutrition and micronutrient malnutrition, largely caused by inadequate access to diverse foods and healthy diets[1–6]. Higher levels of dietary diversity are positively associated with nutrient adequacy, physical and cognitive development, and overall health[7–12]. Households in rural Africa are particularly affected by poor diets and nutritional deficiencies[1,6,13,14]. Many of these rural households depend on farming for their livelihoods. Increasing the diversity of own farm production may help to improve access to nutritious foods and dietary diversity, but the evidence is mixed[15,16].

The question regarding to what extent farm production diversity (FPD) influences the dietary diversity and nutrition of households and individuals has been analysed extensively in recent years, with data from many countries in Africa, Asia and Latin America[15–26]. The results suggest positive associations in many but not all situations. Furthermore, the magnitudes of the associations tend to be small on average, meaning that farm diversification strategies might have limited potential to improve household diets and nutrition at scale[16]. More knowledge is required regarding the conditions needed for farm diversification to be effective and what other strategies may be helpful.

One drawback of the existing work on links between production diversity and dietary diversity is that most studies analyse the situation in one country or setting, so drawing broader international conclusions is difficult. A second drawback is that most studies use cross-section observational data, meaning that unobserved heterogeneity may bias the results. A third drawback is that almost all available studies analyse food production diversity only at the level of individual farms, whereas production diversity at the village level or at higher spatial scales may also influence household dietary diversity. If local markets exist and farms differ in their production patterns, not every single farm would need to be very diverse for households to have access to diverse nutritious foods. There is only one study that analyses production diversity at a higher spatial scale, namely, village-level production diversity and its association with child diets and nutrition[27]. This study uses data from various countries in Africa and evaluates village-level production diversity with remotely sensed crop vegetation data[27]. While this is an interesting approach, the production of fruits, vegetables and animal-sourced foods—all potentially important elements of diversified diets—is not captured in the remotely sensed crop vegetation data.

[1]Center for Development Research (ZEF), University of Bonn, Bonn, Germany. [2]Institute for Food and Resource Economics (ILR), University of Bonn, Bonn, Germany. ✉e-mail: tnguyen3@uni-bonn.de

Table 1 | Household dietary diversity and FPD

	All countries	Ethiopia	Malawi	Niger	Nigeria	Tanzania	Uganda
Household dietary diversity (all households)							
HDDS	5.66 (1.75)	4.42 (1.56)	6.24 (1.75)	5.70 (1.67)	5.98 (1.73)	5.91 (1.63)	5.67 (1.67)
HDDS from own production	1.67 (1.62)	1.39 (1.32)	1.70 (1.58)	0.84 (1.06)	1.23 (1.26)	1.69 (1.72)	2.48 (1.79)
HDDS from market purchases	4.50 (2.23)	3.08 (1.97)	5.00 (2.31)	5.27 (1.92)	5.24 (2.00)	4.62 (2.25)	3.98 (2.11)
HDDS from other sources	0.64 (1.02)	0.27 (0.64)	1.22 (1.25)	0.49 (0.91)	0.55 (1.01)	0.65 (0.97)	0.68 (1.02)
Food from own production (percentage of value)	33.14 (33.52)	42.84 (39.47)	31.11 (31.39)	19.48 (27.34)	23.91 (28.44)	31.72 (34.48)	42.58 (32.12)
Food from market purchases (percentage of value)	59.13 (34.64)	50.02 (39.56)	55.89 (33.56)	75.22 (29.72)	69.79 (29.97)	59.78 (35.89)	50.25 (32.11)
Food from other sources (percentage of value)	7.73 (16.85)	7.14 (19.59)	13.00 (20.42)	5.30 (14.37)	6.30 (14.77)	8.49 (16.94)	7.17 (15.49)
Observations	89,742	13,511	9,163	7,046	18,592	21,117	20,313
FPD (households with own farm production)							
FPD (species)	5.55 (3.33)	7.38 (4.23)	4.04 (2.33)	4.45 (2.37)	4.33 (2.27)	5.85 (3.81)	6.12 (2.78)
Livestock production diversity (food groups)	3.29 (1.51)	3.84 (1.81)	3.08 (1.42)	2.48 (1.09)	2.59 (1.12)	3.55 (1.64)	3.62 (1.28)
Crop production diversity (food groups)	1.07 (0.88)	1.64 (0.96)	0.76 (0.77)	0.98 (0.64)	0.74 (0.60)	1.18 (1.00)	1.05 (0.81)
FPD (food groups)	2.22 (1.14)	2.20 (1.41)	2.32 (1.09)	1.50 (0.99)	1.85 (0.94)	2.38 (1.21)	2.58 (0.91)
Farms growing non-food cash crops (%)	25.63 (43.66)	47.33 (49.93)	18.59 (38.90)	1.05 (10.21)	6.46 (24.58)	30.19 (45.91)	34.28 (47.47)
Observations	67,221	9,981	7,359	5,312	12,882	14,921	16,766

Mean values are shown with standard deviations in parentheses.

Here we address these limitations by using representative panel data with over 89,000 household observations from six African countries—Ethiopia, Malawi, Niger, Nigeria, Tanzania and Uganda. This large and representative sample enhances the external validity of the findings and also allows us to make interesting comparisons. The data stem from the World Bank's Living Standards Measurement Study—Integrated Surveys on Agriculture (LSMS-ISA). They were collected between 2008 and 2022 and include details of agricultural production, food consumption and other relevant socio-economic characteristics. The LSMS-ISA data do not include individual-level dietary data, meaning that our evaluation of dietary diversity remains at the household level. Production diversity is measured in terms of the number of different crop and livestock species as well as the number of food groups produced. We use panel data regression models to estimate how FPD is associated with household dietary diversity, controlling for confounding factors and time-invariant unobserved heterogeneity (Methods). We also examine the role of markets, asset ownership and food production diversity at higher spatial scales for household dietary diversity.

Results

Table 1 presents descriptive statistics of household dietary diversity and FPD in the six African countries. The household dietary diversity score (HDDS) counts the number of different food groups consumed by each household over a period of 7 days. A higher HDDS indicates higher dietary diversity and dietary quality. We see some differences between the countries. The highest mean HDDS is observed in Malawi and the lowest, in Ethiopia.

Across the six countries, farms produce around 5.5 different crop and livestock species on average, which is quite high given the small farm sizes. Over 75% of the farms in the sample have a landholding of less than 2 ha. Strikingly, the average number of species produced is highest in Ethiopia and lowest in Malawi—the exact opposite of what we found for household dietary diversity (Table 1). This is an indication that FPD may not be the most important driver of household dietary diversity.

Table 1 also shows that the number of food groups produced by each farm is lower than the number of food groups consumed by the farm households. Households consume significantly more food groups from market purchases than from own production, in spite of the fact that we exclude foods with low nutrient density—such as sugar, oils and fats, condiments and beverages, which are mainly purchased—from the food group calculations. Also, in terms of overall food consumption, market purchases often matter more than own production. Across the six countries, households obtain 33% of the food values consumed from subsistence production.

Role of FPD in dietary diversity

We now estimate the associations between FPD and the HDDS with regression models, controlling for possible confounding factors (Methods). We find positive and statistically significant estimates in all cases, but the coefficient sizes are relatively small on average (Fig. 1). When FPD is measured in terms of the number of species produced and data from all six countries are included, the mean coefficient is 0.044 implying that households would have to produce over 20 additional species on their farms to increase the HDDS by one food group. The coefficient sizes are larger when FPD is measured in terms of the number of food groups produced. Here the mean coefficient estimate of 0.10 across the six countries implies that households would have to produce 10 additional food groups to increase the HDDS by one unit. However, the coefficient sizes vary by country. The largest coefficient is observed in Niger and the smallest, in Ethiopia. In the following analysis, we measure FPD only in terms of the number of food groups produced.
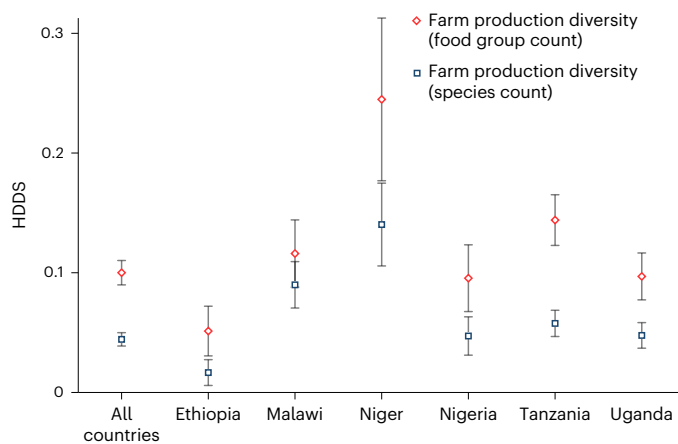
Fig. 1 | Associations between FPD and household dietary diversity. Coefficient estimates from correlated random-effect models are shown with error bars representing 95% confidence intervals. The models with data from all countries include 89,742 observations. Full model results and control variables included are shown in Supplementary Tables 1 and 2.

Other factors influencing dietary diversity

FPD is not the only factor influencing household dietary diversity. The results in Fig. 2 show that several other socio-economic characteristics matter in the context of rural Africa (a breakdown by country is shown in Supplementary Table 2). Weather shocks—defined as the occurrence of a drought, flood, hurricane or related extreme event over the last 12 months—are negatively associated with the HDDS. Several other socio-economic characteristics are positively associated with the HDDS. The production of non-food cash crops—such as cotton, coffee, tea or tobacco—on own farms seems to contribute to higher household dietary diversity through positive cash income effects. Similarly, off-farm wage employment and self-employment in own non-farm enterprises are positively associated with the HDDS. Education of the household head, measured in terms of literacy, also has a large positive estimation coefficient. These findings are in line with earlier studies on links between socio-economic factors, food security and dietary quality[28–36].

Asset ownership is positively associated with dietary diversity. Assets such as mobile phones, motorbikes and electricity are commonly used proxies of wealth in the context of rural Africa. In addition, these assets are directly linked to food market access as well as food preparation and storage, so positive associations with dietary diversity are plausible. Finally, female-headed households have higher HDDSs than male-headed households (Fig. 2). Female-headed households are often worse off in terms of wealth and access to land and other productive resources[37], which might lead to lower welfare and dietary quality. However, in our models, we control for such other factors, which probably explains why we see this positive association here.

FPD and the role of markets

As seen in Fig. 1, the size of the association between FPD and the HDDS varies between countries. However, the association also varies within countries, depending on infrastructure, market access and possibly other spatial conditions. Distance to urban centres—a proxy for the costs of accessing markets in rural Africa[38]—has a negative association with the HDDS in all six study countries (Supplementary Table 3). In the pooled sample, the average household is located around 31 km away from the next urban centre. In addition, as shown in Fig. 3a, with increasing distance to urban centres, the size of the association between FPD and the HDDS rises. In other words, in remote locations with poor market access, FPD plays a more important role for household diets than in better-connected locations. This is expected given that households in remote areas are often more subsistence oriented.

Figure 3b shows that FPD in subsistence households (those with food consumption shares from own production above 50%) is significantly higher than in non-subsistence households. However, at the same time, the HDDS in subsistence households is significantly lower than in non-subsistence households. This comparison suggests that increasing FPD cannot fully compensate for the negative effect of lower market access on household dietary diversity.

Figure 3c shows that there is also a direct link between FPD and subsistence orientation. FPD seems to encourage stronger subsistence orientation, as the association with HDDS from own production is larger than the association with the HDDS from market purchases. In fact, the association between FPD and the HDDS from market purchases is zero or near zero in all countries, except for Niger. The larger coefficient estimate in Niger suggests that farm diversification there is driven not only by subsistence demand but also by market incentives.

Local and regional food production diversity

So far, we have looked at the associations between own FPD and the HDDS. Now we want to understand whether local and regional food production diversity—beyond own farms—may also influence household diets and nutrition. Figure 4 summarizes results from various regressions, comparing the coefficient estimates for farm-level production diversity with those for food production diversity at higher spatial scales, namely, village level, town level and district level (Methods). In the models with data from all countries combined, all local and regional metrics of food production diversity have positive and significant associations with HDDS. Also, in the individual country models, most of the local and regional production diversity metrics have positive associations with household dietary diversity. This is plausible because more diverse agricultural production at the local and regional levels also means higher levels of food diversity available in local and regional markets. Above, we have shown that market purchases are more important for the HDDS than own household production.

Comparing production diversity at different spatial scales, we have found that farm-level production diversity has the largest positive associations with the HDDS in most cases, followed by village level and then town and district levels. This is as expected. Strikingly, however, in Ethiopia and Nigeria, the coefficient sizes of district-level production diversity are at par with those of farm-level production diversity or even larger, underlining the crucial role of local and regional markets.
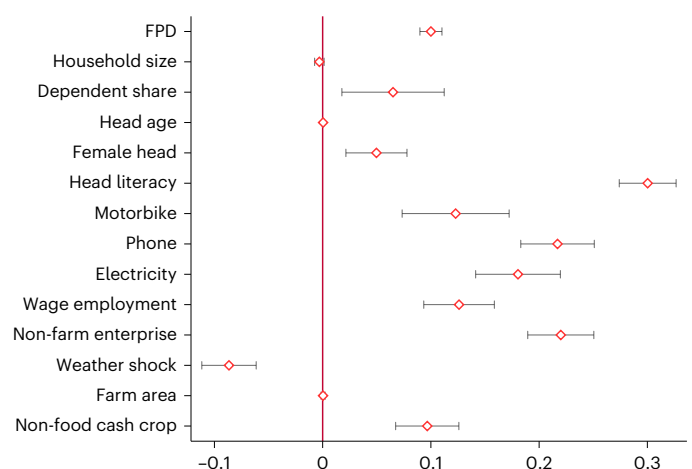

Fig. 2 | Associations between socio-economic variables and household dietary diversity. The dependent variable is household dietary diversity measured in terms of the HDDS. Coefficient estimates from correlated random-effect models are shown with error bars representing 95% confidence intervals (89,742 observations). Full model results and control variables are shown in Supplementary Table 2. FPD measured in terms of the number of food groups produced.
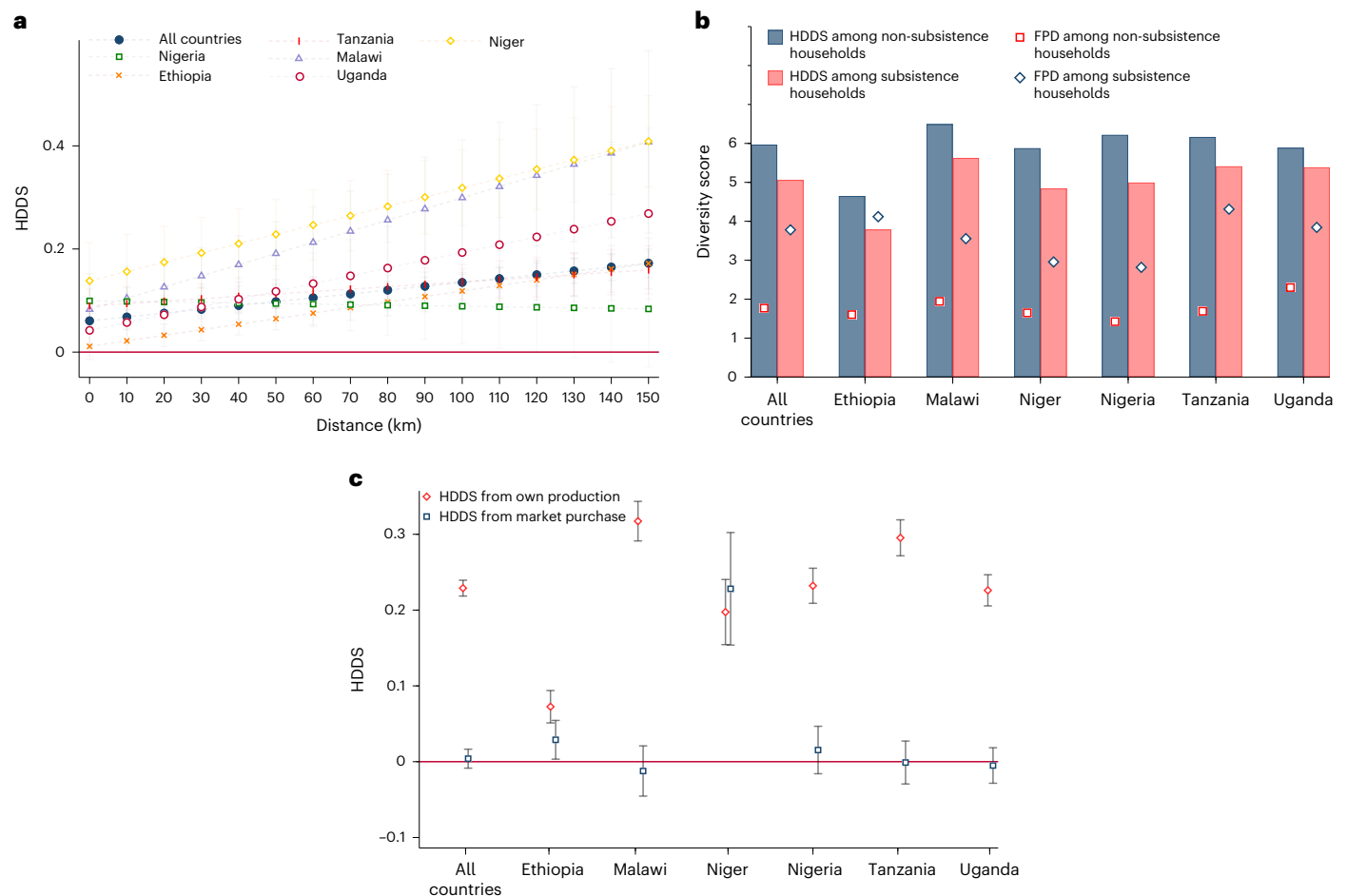
Fig. 3 | Associations between FPD, markets and household diets.
a, Associations between FPD and household dietary diversity by distance to the nearest urban centre. Marginal effects from correlated random-effect models are shown with error bars representing 95% confidence intervals. The model with data from all countries includes 61,916 observations. Full model results and control variables included are shown in Supplementary Table 3. **b,** FPD and household dietary diversity in subsistence and non-subsistence households. Subsistence households are defined as those that obtain more than 50% of their food consumption value from own production. The sample with data from all countries includes 78,003 observations. **c**, Associations between FPD and household dietary diversity obtained from own production and market purchases. Coefficient estimates from correlated random-effect models are shown with error bars representing the 95% confidence intervals. The models with data from all countries include 82,550 observations. Full model results and control variables included are shown in Supplementary Tables 4 and 5. FPD was measured in terms of the number of food groups produced.

Discussion

Our study adds to the existing literature on the links between FPD and household diets and nutrition[15–25] in that we analyse associations with representative panel data from six countries in Africa, namely, Ethiopia, Malawi, Niger, Nigeria, Tanzania and Uganda. We also examine more specifically than earlier work on what factors the size of the associations depend and what roles market access and food production diversity at higher spatial scales may play. Our insights contribute to a better understanding of the complex relationships and can help design strategies for improving rural diets and other important sustainability outcomes. Food system diversification is generally considered a useful approach to enhance sustainability and resilience, but how exactly diversification should look like in specific contexts remains an open question[39,40].

Our results suggest that FPD is positively associated with HDDS. The size of the association depends on how exactly FPD is measured: it is larger when FPD is measured in terms of the number of food groups produced than when it is measured in terms of the number of crop and livestock species produced. This is plausible, as producing and consuming additional species do not necessarily add to the HDDS if the species belong to the same food group (for example, maize, millet, teff, wheat and sorghum all belonging to the food group 'cereals').

However, regardless of how FPD is measured, its mean association with HDDS is relatively small. This is consistent with earlier research[16]. Farms would have to produce many more species and food groups to have a sizeable influence on dietary diversity through this channel. The potentials for drastic farm diversification are limited because most farms in Africa are small and quite diverse anyway. Our results also suggest that farm diversification is often associated with higher levels of subsistence, meaning that very diverse farms may forego development potentials emerging from more market interactions. Markets are important sources of nutritious foods. On average, rural households in all six African countries obtain most food groups from market purchases and only a small fraction from own production. Also, market-oriented households tend to have higher HDDSs than their more subsistence-oriented counterparts.

In spite of small mean coefficient estimates, we find some interesting differences between countries. In Niger, the association between FPD and the HDDS is larger than in the other five countries. Niger is much drier than the other countries and has lower agricultural production potential[41]. Farm diversity in Niger is very low on average; most of the foods are purchased from the market (Table 1). Many farmers in Niger grow cereals with very low average yields. Keeping cattle and small ruminants in pastoral and semi-pastoral systems is also common.
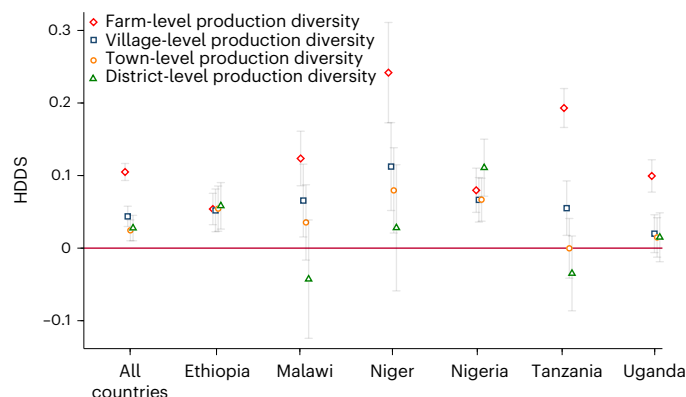
Fig. 4 | Associations between production diversity at different spatial scales and household dietary diversity. Production diversity is measured in terms of the number of food groups produced at each spatial scale. Coefficient estimates from correlated random-effect models are shown with error bars representing 95% confidence intervals. The models with data from all countries include 65,579 observations. Full model results and control variables included are shown in Supplementary Tables 6–9.

Higher production diversity in this context often means that additional livestock species—such as chicken—are kept. Moreover, a few farmers in Niger with access to irrigation grow vegetables and fruits primarily for market sales. Hence, unlike in other countries, in Niger, higher FPD is not associated with subsistence, but with better access to markets and technologies and higher levels of commercialization. This is supported by our finding that FPD in Niger is strongly associated with the HDDS from market purchases (Fig. 3c), which is not the case in the other five countries. These comparisons underline the fact that the effects of FPD depend on the context and that market-led diversification strategies have stronger positive diet and nutrition effects than subsistence-led diversification strategies[16,23,24].

We also find differences within countries. FPD has larger positive associations with the HDDS in remote rural areas than in settings that are better connected to markets and urban centres. This means that farm diversification strategies may be more effective in improving diets in remote locations. Interestingly, however, we find that village-, town- and district-level food production diversity have positive associations with the HDDS as well. In some cases, these associations are in a similar magnitude as those between farm-level production diversity and the HDDS. More diverse local and regional production means that more diverse foods are available and accessible through local and regional markets. These results suggest that diversification strategies should focus not only on individual farms but also on local and regional food systems more broadly.

In settings with good infrastructure conditions and spatially integrated markets, the diversity of foods sold in local markets may not depend much on the diversity of foods produced locally because foods from other regions can easily be imported. However, this is not the case in many rural regions of Africa. With poor infrastructure conditions, perishable nutritious foods—such as vegetables, fruits and dairy—are not traded much into remote rural regions, unless they are produced locally. Hence, local production diversity is important but does not necessarily mean that every individual farm needs to be extremely diverse. For individual farms, high levels of production diversity may face land and labour constraints, hampering efficiency and possible gains from specialization. Local and regional food production diversity can also be achieved if individual farmers specialize to some extent, as long as local markets exist and not all farmers specialize on the same types of output. Strengthening local markets and supply chains for perishable foods could also create new market incentives for farm diversification, such that diversification would

be driven by market opportunities and not primarily by household subsistence needs.

In conclusion, increasing the production diversity of individual farms may be useful in very specific situations but should not be seen as a general strategy to improve dietary quality in rural Africa. In all six countries, market purchases are more important for dietary diversity than own production. Improving market access and the functioning of local markets for nutritious foods should therefore receive higher policy priority. Local and regional production diversity matter, but these are not the same thing as production diversity on individual farms. The appropriate spatial scale needs to be considered when designing food system diversification strategies. Beyond production diversity and food market functioning, access to productivity-enhancing agricultural technologies, infrastructure and non-farm income sources are other important avenues to improve rural diets in Africa, as our results also show.

We acknowledge that there are a few limitations of the data and the approaches used in this study. First, while we use panel data models that allow us to better control for confounding factors than most previous studies, we cannot fully rule out endogeneity bias, so we interpret our estimates as associations, not as causal effects. Second, the nationally representative LSMS-ISA data provide details on household-level food consumption, not on individual-level food intakes. We rely on HDDSs as simple proxies of dietary quality, not considering food quantities, which are also important for nutritional outcomes. Third, links between agricultural production and household food consumption tend to vary seasonally, which is not properly captured in the data. Hence, follow-up studies will be useful to further validate our findings.

Methods
Data
We use data from the LSMS-ISA, focusing on six African countries, namely, Ethiopia, Malawi, Niger, Nigeria, Tanzania and Uganda. The LSMS-ISA surveys are nationally representative and conducted by the country statistical offices with support from the World Bank. In all six countries, a minimum of two survey waves were conducted, using a panel structure. In Ethiopia, data from three waves conducted between 2011 and 2016 are available. In Malawi, four waves (2010–2020) are available; in Niger, two waves (2011–2015); in Nigeria, four waves (2010–2019); in Tanzania, five waves (2008–2022); and in Uganda, seven waves (2009–2020) (Supplementary Table 10). LSMS-ISA surveys were also conducted in Burkina Faso and Mali, but in those two countries, no panel data are available, so we decided to only focus on six countries.

In all countries, the data were collected through face-to-face interviews using structured questionnaires, capturing information on farming and other household economic activities, food and non-food consumption, asset ownership and various other socio-economic factors. Food consumption was captured through 7 day recalls at the household level. In some of the countries, certain data were captured more than once per year. In those cases, we arranged the data in a way that makes them comparable to those of other countries as much as possible. All survey waves from the six countries combined include around 90,000 observations. We exclude observations with missing values or outliers for food consumption and other key variables. Our final sample includes 89,742 observations (Supplementary Table 10).

Not all of the households in the sample are involved in farming. In the main analysis, we include all households, as the decision whether or not to farm is endogenous and may change over time. However, in a robustness check, we also run all regressions only for the subsample of households involved in farming. Furthermore, as sample attrition over the various survey waves occurs, we also run all regressions only with the balanced sample and with the full farmer sample but correcting for possible attrition bias. The results of these robustness checks are very similar to the main results and support the same conclusions (Supplementary Tables 11–18).

Key variables

Dietary diversity. Dietary diversity at the household level is measured in terms of the HDDS, which reflects the economic ability of a household to consume a range of food options and their food security status[42,43]. The HDDS typically includes 12 food groups, namely (1) cereals; (2) legumes, nuts and seeds; (3) vegetables; (4) white roots and tubers; (5) fruits; (6) eggs; (7) milk and related products; (8) fish; (9) meat; (10) spices, condiments and beverages; (11) sugar and sweets; and (12) oils and fats. However, the last three food groups are commonly perceived as less nutritious and, if included, might hinder the capacity of the HDDS to properly reflect dietary quality[21,24,43]. We therefore do not include these three food groups and calculate the HDDS only with the remaining nine groups. We calculate the HDDS based on the 7 day food consumption data collected in the LSMS-ISA surveys.

FPD. FPD is measured by two indicators, namely (1) the number of species produced by each farm and (2) the number of food groups produced by each farm. The species count is a common and intuitive way of measuring the diversity of farms[15,16,21]. However, producing a larger number of species does not necessarily increase food group diversity, because many relevant species produced by farms belong to the same food group[44]. Maize, sorghum and millet are all cereals. Tomatoes, cabbages and eggplants are all vegetables, just to name a few examples. Hence, recent studies have measured FPD also in terms of the number of different food groups produced[19,23,24,45]. For the calculation, we use the same nine food groups as listed above for the HDDS. We use FPD measured in terms of food groups produced for most of the analysis, but provide an explicit comparison of both metrics' associations in Fig. 1.

Local and regional food production diversity. Food production diversity at the village, town and district levels is measured by counting the number of different food groups produced by farms sampled in the respective villages, towns and districts. Note that the administrative levels have different names across the six study countries. Villages are sometimes called clusters or parishes; towns are sometimes called wards or sub-counties; districts are sometimes called departments, counties or woredas. We decided to use the same terminology across study countries to facilitate comparisons. For some of the observations, exact village classifications were missing. These observations were left out for the analysis in Fig. 4. Furthermore, villages with only one sampled farm household were omitted for this part of the analysis because, in these cases, farm-level and village-level production diversity would be identical. On average, the number of farm households sampled per village is seven across the six countries. Of course, the number of food groups produced by sampled households is only a lower-bound estimate of the total food production diversity at the village level because non-sampled households may possibly produce additional food groups. This is a drawback of our approach. However, an advantage of our approach is that we also cover species and food groups that are difficult to evaluate with remotely sensed data[27], including fruits, vegetables and animal-sourced foods.

Regression models

Associations between FPD and the HDDS. We estimate associations between FPD and the HDDS with panel data regression models of the following type:

$$\text{HDDS}_{it} = \alpha + \beta \text{FPD}_{it} + \delta\, \mathbf{X}_{it} + \vartheta \overline{\text{FPD}}_i + + \gamma \overline{X}_i + \theta \mathbf{C}_i + \tau \mathbf{T}_t + \varepsilon_{it} \quad (1)$$

where subscript i indicates households and subscript t indicates survey years. $\mathbf{X}$ is a vector of control variables including asset ownership (farmland, motorbikes and so on), other household farm and socio-economic characteristics (cash crop production, off-farm employment, access to infrastructure and institutions, and so on), characteristics of the

household head (sex, age, education) and weather shocks (see Supplementary Table 19 for a full list and descriptive statistics). $\mathbf{C}$ and $\mathbf{T}$ are vectors of country and time fixed effects, the former of which is only included in the pooled models with data from all six countries. $\alpha, \beta, \delta, \upsilon, \gamma, \theta$ and τ are parameters to be estimated, and ε is a random error term. We are particularly interested in the estimate of the coefficient β. A positive and statistically significant β would indicate that FPD is positively associated with the HDDS.

One concern in estimating such models with observational data is the possible endogeneity of FPD. FPD is probably influenced by many observed and unobserved factors. While controlling for observed heterogeneity is straightforward, controlling for unobserved heterogeneity is sometimes not. The advantage of panel data over cross-sectional data is that we can control for time-invariant unobserved heterogeneity, either through including household fixed effects (FE) or correlated random effects (CRE)[46]. Both the FE and the CRE estimators do not inherently control for time-varying heterogeneity. The FE estimator fully relies on variation within households over time, so coefficients for explanatory variables with limited variation over time can sometimes not be estimated efficiently and time-invariant variables are dropped. As we are also interested in the coefficients of variables with limited variation over time, we use the CRE estimator for the main analysis and include potentially endogenous variables as time averages, indicated by $\overline{\text{FPD}}$ and $\overline{X}$ in equation (1). However, in robustness checks, we also use the FE estimator. In addition, as the HDDS is a count variable, we conduct additional robustness checks with Poisson specifications. Results from the FE and Poisson CRE specifications are shown in Supplementary Tables 20–27. They are similar to the original estimates and support the same conclusions.

In addition to the standard models in equation (1), we also analyse how FPD is associated with dietary diversity from different sources by disaggregating HDDS into (1) HDDS from market purchases and (2) HDDS from own production, using both as outcome variables in separate regressions. Furthermore, to analyse the role of market access and urban proximity, in some of the regressions, we include distance to the nearest urban centre (defined as a city with a population of at least 20,000) and an interaction term between FPD and this distance variable as additional regressors.

Role of local and regional production diversity for the HDDS. We estimate associations between local and regional production diversity and the HDDS with regression models of the following type:

$$\text{HDDS}_{ivt} = \omega + \pi \text{LPD}_{vt} + \varphi \mathbf{X}_{ivt} + \sigma \overline{\text{LPD}}_v + \rho \overline{X}_{iv} + \psi N_{vt} + \varsigma \mathbf{C}_i + \Omega \mathbf{T}_t + \varepsilon_{ivt} \quad (2)$$

where LPD_{vt} stands for local (or regional) production diversity at the spatial scale of interest v, which can be the village, the town or the district in which household i resides. As the number of farmers sampled at each spatial scale differs, we additionally control for the number of sampled farmers, N_{vt}. $\omega, \pi, \varphi, \sigma, \rho, \psi, \varsigma$ and Ω are parameters to be estimated. All other variables in equation (2) are defined as in equation (1).

As explained above, for this part of the analysis, we had to exclude some of the observations. The total number of observations used for estimating equation (2) is 65,579 in the pooled country model. To compare the estimates for FPD and local and regional production diversity at different spatial scales, we use this same smaller sample for all estimates shown in Fig. 4. The coefficients for FPD with this smaller sample hardly differ from the full-sample estimates in Fig. 1, underlining the overall robustness of our findings.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The data used for this analysis are available at https://www.worldbank.org/en/programs/lsms/initiatives/lsms-ISA.

Code availability

The code used for this analysis is available at https://doi.org/10.7910/DVN/NYPA0Z.

References

1. Micha, R. et al. *2022 Global Nutrition Report: Stronger Commitments for Greater Action* (Development Initiatives, 2022); https://globalnutritionreport.org/documents/922/2022_Global_Nutrition_Report.pdf
2. FAO, IFAD, UNICEF, WFP & WHO. *The State of Food Security and Nutrition in the World 2024* (FAO, 2024); https://doi.org/10.4060/cd1254en
3. Stevens, G. A. et al. Micronutrient deficiencies among preschool-aged children and women of reproductive age worldwide: a pooled analysis of individual-level data from population-representative surveys. *Lancet Glob. Health* **10**, e1590–e1599 (2022).
4. Wang, X. et al. Global food nutrients analysis reveals alarming gaps and daunting challenges. *Nat. Food* **4**, 1007–1017 (2023).
5. Miller, V. et al. Global dietary quality in 185 countries from 1990 to 2018 show wide differences by nation, age, education, and urbanicity. *Nat. Food* **3**, 694–702 (2022).
6. McCullough, E. B., Lu, M., Nouve, Y., Arsenault, J. & Zhen, C. Nutrient adequacy for poor households in Africa would improve with higher income but not necessarily with lower food prices. *Nat. Food* **5**, 171–181 (2024).
7. Ruel, M. T. Operationalizing dietary diversity: a review of measurement issues and research priorities. *J. Nutr.* **133**, 3911S–3926S (2003).
8. Lachat, C. et al. Dietary species richness as a measure of food biodiversity and nutritional quality of diets. *Proc. Natl Acad. Sci. USA* **115**, 127–132 (2018).
9. Mozaffarian, D., Rosenberg, I. & Uauy, R. History of modern nutrition science—implications for current research, dietary guidelines, and food policy. *BMJ* **361**, k2392 (2018).
10. Xiao, C. et al. Associations of dietary diversity with the gut microbiome, fecal metabolites, and host metabolism: results from 2 prospective Chinese cohorts. *Am. J. Clin. Nutr.* **116**, 1049–1058 (2022).
11. Korir, L., Rizov, M. & Ruto, E. Diet diversity, malnutrition and health: evidence from Kenya. *J. Agric. Econ.* **74**, 534–550 (2023).
12. Liu, D. et al. Association of dietary diversity changes and mortality among older people: a prospective cohort study. *Clin. Nutr.* **40**, 2620–2629 (2021).
13. Headey, D., Hirvonen, K., Hoddinott, J. & Stifel, D. Rural food markets and child nutrition. *Am. J. Agric. Econ.* **101**, 1311–1327 (2019).
14. Sibhatu, K. T. & Qaim, M. Rural food security, subsistence agriculture, and seasonality. *PLoS ONE* **12**, e0186406 (2017).
15. Jones, A. D. Critical review of the emerging research evidence on agricultural biodiversity, diet diversity, and nutritional status in low- and middle-income countries. *Nutr. Rev.* **75**, 769–782 (2017).
16. Sibhatu, K. T. & Qaim, M. Meta-analysis of the association between production diversity, diets, and nutrition in smallholder farm households. *Food Policy* **77**, 1–18 (2018).
17. Khonje, M. G., Ricker-Gilbert, J., Muyanga, M. & Qaim, M. Farm-level production diversity and child and adolescent nutrition in rural sub-Saharan Africa: a multicountry, longitudinal study. *Lancet Planet. Health* **6**, e391–e399 (2022).
18. Fanzo, J. C. Decisive decisions on production compared with market strategies to improve diets in rural Africa. *J. Nutr.* **147**, 1–2 (2017).
19. Hirvonen, K. & Hoddinott, J. Agricultural production and children's diets: evidence from rural Ethiopia. *Agric. Econ.* **48**, 469–480 (2017).
20. Ruel, M. et al. Nutrition-sensitive agriculture: what have we learned so far? *Glob. Food Secur.* **17**, 128–153 (2018).
21. Sibhatu, K. T., Krishna, V. V. & Qaim, M. Production diversity and dietary diversity in smallholder farm households. *Proc. Natl Acad. Sci. USA* **112**, 10657–10662 (2015).
22. Zanello, G., Shankar, B. & Poole, N. Buy or make? Agricultural production diversity, markets and dietary diversity in Afghanistan. *Food Policy* **87**, 101731 (2019).
23. Muthini, D., Nzuma, J. & Qaim, M. Subsistence production, markets, and dietary diversity in the Kenyan small farm sector. *Food Policy* **97**, 101956 (2020).
24. Koppmair, S., Kassie, M. & Qaim, M. Farm production, market access and dietary diversity in Malawi. *Public Health Nutr.* **20**, 325–335 (2017).
25. Ecker, O. Agricultural transformation and food and nutrition security in Ghana: does farm production diversity (still) matter for household dietary diversity? *Food Policy* **79**, 271–282 (2018).
26. Rasmussen, L. V. et al. Joint environmental and social benefits from diversified agriculture. *Science* **384**, 87–93 (2024).
27. Tobin, D., Jones, K. & Thiede, B. C. Does crop diversity at the village level influence child nutrition security? Evidence from 11 sub-Saharan African countries. *Popul. Environ.* **41**, 74–97 (2019).
28. Nguyen, T. T., Nguyen, T. T., Le, V. H., Managi, S. & Grote, U. Reported weather shocks and rural household welfare: evidence from panel data in northeast Thailand and central Vietnam. *Weather Clim. Extremes* **30**, 100286 (2020).
29. Akerele, D., Momoh, S., Aromolaran, A. B., Oguntona, C. R. & Shittu, A. M. Food insecurity and coping strategies in South-West Nigeria. *Food Secur.* **5**, 407–414 (2013).
30. Feleke, S. T., Kilmer, R. L. & Gladwin, C. H. Determinants of food security in southern Ethiopia at the household level. *Agric. Econ.* **33**, 351–363 (2005).
31. Babatunde, R. O. & Qaim, M. Impact of off-farm income on food security and nutrition in Nigeria. *Food Policy* **35**, 303–311 (2010).
32. Do, T. L., Nguyen, T. T. & Grote, U. Nonfarm employment and household food security: evidence from panel data for rural Cambodia. *Food Secur.* **11**, 703–718 (2019).
33. Rahman, A. & Mishra, S. Does non-farm income affect food security? Evidence from India. *J. Dev. Stud.* **56**, 1190–1209 (2020).
34. Parlasca, M. C., Mußhoff, O. & Qaim, M. Can mobile phones improve nutrition among pastoral communities? Panel data evidence from northern Kenya. *Agric. Econ.* **51**, 475–488 (2020).
35. Nguyen, T. T., Nguyen, T. T., Do, M. H., Nguyen, D. L. & Grote, U. Shocks, agricultural productivity, and natural resource extraction in rural Southeast Asia. *World Dev.* **159**, 106043 (2022).
36. Candelise, C., Saccone, D. & Vallino, E. An empirical assessment of the effects of electricity access on food security. *World Dev.* **141**, 105390 (2021).
37. Quisumbing, A. R. & Doss, C. in *Handbook of Agricultural Economics* Vol. 5 (eds Barrett, C. B. & Just, D. R.) 4481–4549 (North-Holland, 2021).
38. Hülsen, V., Khonje, M. G. & Qaim, M. Market food environments and child nutrition. *Food Policy* **128**, 102704 (2024).
39. Dwivedi, S. L. et al. Diversifying food systems in the pursuit of sustainable food production and healthy diets. *Trends Plant Sci.* **22**, 842–856 (2017).
40. Hertel, T., Elouafi, I., Tanticharoen, M. & Ewert, F. Diversification for enhanced food systems resilience. *Nat. Food* **2**, 832–834 (2021).

41. Dzanku, F. M., Liverpool-Tasie, L. S. O. & Reardon, T. The importance and determinants of purchases in rural food consumption in Africa: implications for food security strategies. *Glob. Food Secur.* **40**, 100739 (2024).

42. *Guidelines for Measuring Household and Individual Dietary Diversity* (FAO, 2013); https://www.fao.org/3/i1983e/i1983e00.pdf

43. Verger, E. O., Ballard, T. J., Dop, M. C. & Martin-Prevel, Y. Systematic review of use and interpretation of dietary diversity indicators in nutrition-sensitive agriculture literature. *Glob. Food Sec.* **20**, 156–169 (2019).

44. Berti, P. R. Relationship between production diversity and dietary diversity depends on how number of foods is counted. *Proc. Natl Acad. Sci. USA* **112**, e5656 (2015).

45. Sibhatu, K. T. & Qaim, M. Farm production diversity and dietary quality: linkages and measurement issues. *Food Secur.* **10**, 47–59 (2018).

46. Wooldridge, J. M. Correlated random effects models with unbalanced panels. *J. Econom.* **211**, 137–150 (2019).

Author contributions

T.-T.N. and M.Q. conceptualized and designed the study. T.-T.N. carried out the data analysis. T.-T.N. and M.Q. wrote the paper.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s43016-024-01096-6.

Correspondence and requests for materials should be addressed to Thanh-Tung Nguyen.

Peer review information *Nature Food* thanks Steven Block and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	We only use secondary data in this study. We did not collect data and did not use any software for data collection.
Data analysis	Data analysis is conducted by using the Stata software, version 16.0. https://doi.org/10.7910/DVN/NYPA0Z

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:
- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

The data are publicly available for research purposes from https://www.worldbank.org/en/programs/lsms/initiatives/lsms-ISA

Research involving human participants, their data, or biological material

Policy information about studies with <u>human participants or human data</u>. See also policy information about <u>sex, gender (identity/presentation), and sexual orientation</u> and <u>race, ethnicity and racism</u>.

Reporting on sex and gender	N/A (we use secondary data)
Reporting on race, ethnicity, or other socially relevant groupings	N/A (we use secondary data)
Population characteristics	N/A (we use secondary data)
Recruitment	N/A (we use secondary data)
Ethics oversight	N/A (we use secondary data)

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☐ Life sciences   ☒ Behavioural & social sciences   ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see <u>nature.com/documents/nr-reporting-summary-flat.pdf</u>

Behavioural & social sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description	Study with observational, quantitative data and a longitudidal (panel) design. The data are evaluated with panel data models and correlated random effects (pseudo fixed effects) estimators.
Research sample	The data stem from the World Bank's Living Standards Measurement Study – Integrated Surveys on Agriculture (LSMS-ISA). The samples are nationally representative, including randomly selected households in six African countries. These countries were selected due to the availability of panel datasets and the widespread issues of undernutrition and low dietary quality.
Sampling strategy	We did not collect the data ourselves but use secondary LSMS-ISA data. The data were collected by the official statistical offices in the study countries with support from the World Bank as part of the larger LSMS-ISA initiative. Households were sampled using the stratified random sampling procedure. First, primary sampling units (PSU) within major zone strata selected based on the national population frame, followed by the second stage that households are randomly selected. Detailed information on the sampling procedure in each country is provided in https://www.worldbank.org/en/programs/lsms/initiatives/lsms-ISA
Data collection	We did not collect the data ourselves but use secondary LSMS-ISA data. In all countries, the LSMS-ISA data were collected through face-to-face interviews using structured questionnaires.
Timing	The data in the six study countries were collected in multiple survey waves. In Ethiopia, 3 waves (09/11-03/12, 09/13-04/14, 09/15-04/16), in Malawi, 4 waves (03/10-03/11, 04/13-11/13, 04/16-04/17, 04/19-03/20), in Niger, 2 waves (07/11-01/12, 09/14-03/15), in Nigeria, 4 waves (08/10-04/11; 09/12-04/13, 08/15-04/16; 07/18-02/2019), in Tanzania, 5 waves (10/08-09/09, 10/10-11/11, 10/12-11/13, 10/14-10/15, 01/19-01/22), and in Uganda, 7 waves (09/09-08/10, 11/10-10/11, 11/11-12/12, 09/13-08/14, 03/15-03/16, 03/18-02/19, 03/19-02/20) are available.
Data exclusions	We use the data as publicly provided on the LSMS-ISA website. All survey waves from the six countries combined include around 90,000 observations. We excluded observations with missing values or outliers for food consumption and other key variables. Our final sample includes 89,742 observations
Non-participation	Only households that agreed to be surveyed were included (https://www.worldbank.org/en/programs/lsms/inititiatives/lsms-ISA)
Randomization	Data were collected through stratified random sampling. The data are observational, no interventions were conducted.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒ ☐	Antibodies
☒ ☐	Eukaryotic cell lines
☒ ☐	Palaeontology and archaeology
☒ ☐	Animals and other organisms
☒ ☐	Clinical data
☒ ☐	Dual use research of concern
☒ ☐	Plants

Methods

n/a	Involved in the study
☒ ☐	ChIP-seq
☒ ☐	Flow cytometry
☒ ☐	MRI-based neuroimaging

Plants

Seed stocks	N/A
Novel plant genotypes	N/A
Authentication	N/A