Mapping the genetic and environmental aetiology of autistic traits in Sweden and the United Kingdom

Zoe E. Reed[1,2], Henrik Larsson[4,5], Claire M.A. Haworth[1,3,6,7], Dheeraj Rai[2,3,8], Sebastian Lundström[9], Angelica Ronald[10], Abraham Reichenberg[11], Paul Lichtenstein[12], Oliver S.P. Davis[1,2,3,7]

1) MRC Integrative Epidemiology Unit, University of Bristol, UK

2) Department of Population Health Sciences, Bristol Medical School, University of Bristol, UK

3) National Institute for Health Research Biomedical Research Centre, University Hospitals Bristol NHS Foundation Trust and the University of Bristol, Bristol, UK

4) School of Medical Sciences, Örebro University, Sweden

5) Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Sweden

6) School of Psychological Science, University of Bristol, UK

7) The Alan Turing Institute, London, UK

8) Avon and Wiltshire Partnership NHS Mental Health Trust

9) Gillberg Neuropsychiatry Centre, Centre for Ethics, Law and Mental Health, University of Gothenburg, Sweden

10) Department of Psychological Sciences, Birkbeck, University of London, UK

11) Department of Psychiatry and Seaver Autism Center for Research and Treatment, Ichan School of Medicine at Mount Sinai, New York, USA

12) Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Sweden

Abstract

Background: Autistic traits are influenced by both genetic and environmental factors, and are known to vary geographically in prevalence. But to what extent does their aetiology also vary from place to place?

Methods: We applied a novel spatial approach to data on autistic traits from two large twin studies, the Child and Adolescent Twin Study in Sweden (CATSS; N=16,677, including 8,307 twin pairs) and the Twins Early Development Study in the UK (TEDS; N=11,594, including 5,796 twin pairs), to explore how the influence of nature and nurture on autistic traits varies from place to place.

Results: We present maps of gene- and environment- by geography interactions in Sweden and the United Kingdom (UK). Our results suggest that there is higher heritability and lower non-shared environmental influences on autistic traits in more densely populated areas. For example, we observe greater heritability in the more populated and urban areas in southern Sweden and near the main cities in the UK. Non-shared environmental influences are less consistent across the two countries and although these influences tend to be higher in more rural areas, there are also areas of greater influence around cities.

Conclusions: We hope this systematic approach to aetiological interactions will inspire research to identify previously unknown environmental influences on the aetiology of autistic traits. By doing so, we can gain greater understanding of how these environments draw out or mask genetic predisposition and other

environmental influences and could lead to health and social policy innovations to support those with ASD and autistic traits.

Introduction

Autism spectrum disorder (ASD) is a neurodevelopmental condition, manifesting in childhood and characterised by persistent difficulties with social communication and repetitive behaviours. ASD has a significant impact on child development, often including language difficulties and other co-occurring conditions which may persist into adulthood (Seltzer, Shattuck, Abbeduto, & Greenberg, 2004).

The aetiology of ASD reflects both genetic and environmental influences. Twin and family studies suggest that genetic differences between people explain around 80% of the population variance for ASD (Bai et al., 2019; Sandin et al., 2017; Tick, Bolton, Happé, Rutter, & Rijsdijk, 2016). Most studies suggest that the remaining variance is explained by variation in the non-shared environment. That is, environmental influences that do not contribute to similarity within families.

Reported prevalence of ASD varies, with estimates in developed countries between 1% and 1.5% (Idring et al., 2015; Lundström, Reichenberg, Anckarsäter, Lichtenstein, & Gillberg, 2015; Lyall et al., 2017), although this is known to vary across geographical regions (Bakian, Bilder, Coon, & McMahon, 2015; Campbell, Reynolds, Cunningham, Minnis, & Gillberg, 2011; Chen, Liu, Su, Huang, & Lin, 2008; Hoffman et al., 2017; Mazumdar, Winter, Liu, & Bearman, 2013). Several studies suggest that living in an urban environment is associated with greater risk of ASD compared to rural environments (Chen et al., 2008; Lauritsen et al., 2014; Vassos, Agerbo, Mors, & Bøcker Pedersen, 2016; Wu & Jackson, 2017).

Possible reasons given for this geographical variation in prevalence of ASD include regional diagnostic bias, differences in access to health services or diagnostic resources, different levels of parental awareness, air pollution exposure during pregnancy, green space availability and local trends in socioeconomic status.

If prevalence of autistic traits varies from place to place, is the same true of the aetiology? For example, does variation in the environment explain variation in autistic traits in some areas more than others? Or does the environment in some areas draw out genetic differences between children in their propensity for developing autistic traits? We previously developed a spatial approach to twin model-fitting called spACE to detect spatial variation in genetic and environmental influences within a country (Davis, Haworth, Lewis, & Plomin, 2012). This approach has the potential to highlight gene-environment (G×E) and environment-environment (E×E) interactions for outcomes such as autistic traits. G×E and E×E represent variation in aetiological influences on a trait depending on environmental exposure. For example, genetic risk of a mental health disorder may be drawn out by a stressful environment or genetic risk of hay fever may only reveal itself in pollen-rich areas. The spACE approach allows us to investigate this, mapping geographical patterns of nature and nurture without requiring the measurement of specific genetic variants or specific environmental characteristics.

Here we apply the spACE approach to data on autistic traits in Sweden and the UK. Autistic traits and diagnostic categories of ASD show substantial aetiological overlap (Colvert et al., 2015; Robinson et al., 2016), with genetic correlations

from bivariate twin models of 0.52-0.89. The heritability of autistic traits does not change as a function of severity (Lundström et al., 2012; Robinson et al., 2011; Angelica Ronald et al., 2006), and genetic links have been identified between extreme and sub-threshold variation in ASD (Robinson et al., 2011; Angelica Ronald et al., 2006), so to maximise power we have focussed on trait measures rather than diagnoses.

It seems plausible that environments previously identified as important for the development of autistic traits will also influence aetiology. For example, given previous research on the social stress of urban compared to rural upbringing (Lederbogen et al., 2011), we hypothesise that urban-rural differences will be apparent in the aetiology of autistic traits. However, more importantly, we hope that by systematically mapping geographical differences in aetiology we will facilitate identification of new environments and shed light on the mechanisms by which they act.

Methods and materials

The Swedish Twin Registry and CATSS

CATSS (Anckarsäter et al., 2011), a sub-study of the Swedish twin registry (Magnusson et al., 2012), was launched in 2004 to investigate childhood-onset neurodevelopmental problems such as ADHD and ASD in childhood and adolescence, for all twins turning 9 or 12 years since 2004. Parents were asked to participate in a telephone interview to collect information on various health-related issues. By 2013, when data on autistic traits were obtained, 8,610 parents had responded to this request, accounting for 17,220 twins. The CATSS-9/12 study obtained ethical approval from the Karolinska Institute Ethical Review Board: Dnr 03-672 and 2010/507-31/1, CATSS-9 – clinical 2010/1099-31/3 CATSS-15 Dnr: 2009/1599-32/5, CATSS-15/DOGSS Dnr: 03-672 and 2010/1356/31/1, and CATSS-18 Dnr: 2010/1410/31/1.

For autistic traits, 16,677 participants had data available (including 8,307 complete pairs and 63 incomplete pairs of twins, 51% were male).

CATSS measures of autistic traits

The Autism-Tics, ADHD and other Comorbidities (A-TAC) inventory, based on the Diagnostic and Statistical Manual of Mental Disorders (DSM)-IV criteria, was used in the telephone interview with parents to collect information on a range of neurodevelopmental problems. This inventory has previously been validated in both clinically diagnosed children and the general population (Anckarsäter et al., 2008; Hansson et al., 2005; Larson et al., 2010, 2013; Mårland et al., 2017). The inventory includes 17 items that assess autistic traits, where respondents can answer 'yes/1', 'yes,

to some extent/0.5', and 'no/0'. Following the standard approach, we created a score for each individual by summing these item scores.

As expected, in our sample the median score was 0.00 (interquartile range [IQR]=1.00) for autistic traits. In previous validation studies a low and high cut-off of 4.5 and 8.5 for ASD have been established for broad screening and for use as a clinical proxy, respectively. Most people in this general population sample score well below these cut-offs.

CATSS location data

To conduct the spACE analysis, we assigned a geographical location to each family. In CATSS we matched each twin pair to a Small Areas for Market Statistics (SAMS) location, for the most recent location data we had available up to 2009, using data from Statistics Sweden (http://www.scb.se/en/) and assigned coordinates based on the centroid of the SAMS location.

To provide context for the results for Sweden, it is useful to understand a little about its geography. **Figure 1** shows a map of Sweden and some general indicators of the country's geography; the **supplementary materials** contain a detailed description.

FIGURE 1

The Twins Early Development Study

The Twins Early Development Study (TEDS) contacted parents of twins born in England and Wales between January 1994 and December 1996 (Haworth, Davis, & Plomin,

2013). 16,810 pairs of twins were initially recruited, and currently there are over 10,000 twin pairs still enrolled in TEDS. The participants are demographically representative of the UK population of a similar age, with the majority identifying themselves as white British and with English as their first language. TEDS has collected wide-ranging data on cognitive and behavioural development, using approaches that include questionnaire booklets, telephone testing and web-based tests. The twins, their parents and teachers have all participated in data collection. Ethical approval for TEDS research is provided by the Institute of Psychiatry, Psychology and Neuroscience Ethics Committee, King's College London.

Full phenotypic data for autistic traits were available for 11,594 TEDS participants (including 5,796 complete pairs and 62 incomplete pairs of twins). For these twins the mean age was 11.30 (SD=0.72) and 48% were male.

TEDS measures of autistic traits

Parents in TEDS completed the Childhood Autism Spectrum Test (CAST) when the twins were age 12 years. The CAST consists of 30 items, scored 1 for yes or 0 for no (Scott, Baron-Cohen, Bolton, & Brayne, 2002). For participants included in our analyses, the median score for autistic traits was 4.0 (IQR=4.84). The CAST score considered indicative of ASD is 15.

TEDS location data

We assigned each twin pair geographical coordinates based on the centroids of their postcodes at age 12. To provide context for the results for the UK, **Figure 2** displays a

map and some general indicators of the country's geography; the **supplementary materials** include a detailed description.

FIGURE 2

Statistical analyses

ACE models and maps in CATSS and TEDS

In twin analysis, within-pair similarity of monozygotic (MZ) and dizygotic (DZ) twins is compared to estimate parameters for additive genetic (A), shared environmental (C) and non-shared environmental (E) influences on a trait. In this context, the shared environment refers to influences other than DNA similarity that make children growing up in the same family more similar to each other, whilst the non-shared environment refers to influences that do not contribute to similarity within families. It is not possible to assign specific environments to one or the other environmental component because most environments themselves show both shared and non-shared (and often genetic) influences. We can estimate the contribution of genetic and environmental influences because of the different ways these influences are shared in MZ and DZ twin pairs. For MZ twins, who share 100% of their segregating alleles, A influences correlate 1, whereas for DZ twins they correlate 0.5 because DZ twins share, on average, 50% of their segregating alleles. For both MZ and DZ twins growing up in the same family the shared environmental correlation is 1. In contrast, the non-shared environment is uncorrelated and contributes to differences between twins (Rijsdijk & Sham, 2002).

In this study, we applied a version of the spACE analysis method (Davis et al., 2012) to explore how A, C and E for autistic traits vary geographically. We fit full information

maximum likelihood structural equation models to twin data in R (version 3.3.1) using the OpenMx package (version 2.9.4), calculating A, C and E at many different target locations across an area. In this study we built on our previous work by applying the weights within the structural equation modelling framework, rather than by calculating weighted correlation matrices and using those as input. In twin analysis it is possible to model non-additive genetic effects (D) instead of shared environmental effects (C), and D influences are sometimes found with ASD. However, the D component is highly correlated with the A component, which means confidence intervals are wide and the tendency of variance to swap between these two components makes it difficult to compare results across locations. Because of this, we have fitted ACE models, so here A should be considered broad-sense heritability, including both additive and non-additive genetic influences.

For target locations in Sweden we used the centroid of each unique SAMS that included at least one twin pair. Because UK postcodes give more precise locations than Swedish SAMS, we instead selected UK target locations representative of local population density to preserve participant anonymity. All twin pairs contributed to the results at each location, but contributions were weighted according to the distance of each twin pair from the target location:

$$w_i(x) = \frac{1}{d(x, x_i)^p}$$

where x represents the target location, x_i represents the location of a twin pair, d is the Euclidean distance between x and x_i, and p is the power parameter that controls the rate of drop-off of a twin pair's influence over distance (0.5 for these analyses). We included

sex as a covariate in all the models (accounting for 2.59% of the variance), and age in the TEDS data (accounting for 0.34% of the variance). Further detail on the spACE approach can be found in (Davis et al., 2012). We plotted maps to visualise results (**figures 3 and 4**), where each target location is coloured according to the value of the estimate at that location compared to the mean of values across the map. Low values appear blue and high values appear red, with increasing salience of the colour representing increasing distance from the mean. To avoid outliers having a large effect on the distribution of colours in the maps, we assigned the highest 4% of values to the brightest red and the lowest 4% of values to the brightest blue before assigning colour values to equal ranges between the two. The histograms show the distribution of results and the corresponding colours. Higher resolution and interactive versions of the maps can be found at https://dynamicgenetics.github.io/spACEjs/.

We estimated 95% confidence intervals for A, C and E at each target location and using the CATSS data we performed sensitivity analyses for how A, C and E estimates vary based on the historical residential location used for the twin pairs (**supplementary video 1**). This may allow for identification of critical developmental periods when the geographical environment is particularly influential; for example, if clear patterns are seen when participant locations for the analysis are based on their location at a specific age.

Sex limitation models

While some previous studies have identified no aetiological sex differences for ASD (Constantino & Todd, 2003; Angelica Ronald et al., 2006), others have. For example, modest sex differences were found in previous work with the Swedish Twin Study (A

Ronald, Larsson, Anckarsäter, & Lichtenstein, 2011). To maximise power, we report results that equate the aetiological influences for males and females in the main text. But for the Swedish data, where we replicated small quantitative sex differences in aetiology, we conducted further separate analyses for males and females (**supplementary FigS4 and FigS5**).

Data availability

The data used in this study are available to researchers directly from CATSS and TEDS. Procedures for accessing the data are described at https://ki.se/en/meb/the-child-and-adolescent-twin-study-in-sweden-catss and https://www.teds.ac.uk/researchers/teds-data-access-policy.

Code availability

Code that implements the spACE model and the interactive online maps is available under a GPLv3 open source licence from the scripts directory at https://github.com/DynamicGenetics/spACEjs/.

Results

Mapping the aetiology of autistic traits in Sweden

We mapped the results from each of the 4,199 locations (**Figure 3** and https://dynamicgenetics.github.io/spACEjs/). Because we modelled raw variance after standardising data to mean 0 and SD 1 at the population level (i.e. we did not standardize the A, C and E estimates at each location to add up to one) genetic and environmental influences are not reciprocal at each location, so it is possible for a location to show both strong genetic and environmental influences. Maps with A and E constrained to add up to one in each location (i.e. proportional) are shown in **supplementary FigS2** and show similar results to those for the raw variance. For comparison, we have also plotted results of the weighted means of scaled autistic trait scores at the same locations in **supplementary FigS3** and observe geographical variation for mean autistic trait scores, reflecting the expected variation in prevalence.

FIGURE 3

The results suggest that the amount of variation in autistic traits explained by genetic influences (A) is generally greater in urban areas and less in the sparsely populated north and more rural southern belt. The non-shared environment (E) frequently shows the opposite pattern, with the variation explained generally less in and around the capital and more in southern and northern rural areas. However, we also observe greater contribution of E in the areas around the cities of Gothenburg and Malmö. Variation in A and E can also be seen within local areas, such as around Stockholm, where there are both low and high values for A, suggesting genetic influences are

moderated by factors beyond urbanicity. The histograms for the raw variance indicate that the variance explained by genetic influences ranges from 0.55 to 0.91, with a mean of 0.65 (SD=0.02). The variance explained by E ranges from 0.25 to 0.46, with a mean of 0.34 (SD=0.02). Variation in autistic traits explained by C was approximately zero over the whole of Sweden. Confidence intervals for estimates at each location are provided in **supplementary TableS1. Supplementary video 1** shows that the overall patterns for variation in A and E remained similar irrespective of the historical location used for each twin pair.

Mapping the aetiology of autistic traits in the UK

Figure 4 (and https://dynamicgenetics.github.io/spACEjs/) maps genetic and environmental influences on autistic traits in the UK at 6,758 locations. Again, this is a map of the raw variance. Maps with A, C and E constrained to add to one at each location are shown in **supplementary FigS6**, with very similar results.

FIGURE 4

The raw results for A are consistent with those in Sweden where we observed higher heritability in more densely populated areas. The mean of A is slightly higher in the UK (0.76, SD=0.01) than in Sweden (0.65, SD=0.02). For non-shared environment (E) the patterns are less similar across countries. London, the capital city, and the surrounding south-east of the UK show greater influence of E compared to the north and some regions in the mid-west of England and Wales. In contrast, Sweden's capital, Stockholm, and the surrounding areas show lower estimates of E. Again, C is approximately zero for autistic traits across all regions. As before, local variation in A and E is apparent within

large cities such as London. As the histograms show, A is fairly normally distributed between 0.69 and 0.84 across regions. E ranges more narrowly from 0.21 to 0.29 in a bimodal distribution with a positive skew. Confidence intervals for estimates at each location are provided in **supplementary TableS3.**

Discussion

Our results are consistent with previous population-level estimates of genetic and environmental influences, and demonstrate geographical variation in genetic and non-shared environmental influences on autistic traits in Sweden and the UK.

Geographical differences in genetic influences on autistic traits are indicative of gene-environment interactions where the interacting environmental variable varies by location. For areas of increased genetic influence this means that the environment in these areas draws out genetic influence, in the same way that the presence of airborne pollen would reveal individual differences in genetic risk for hay fever. Areas of increased environmental influence imply regions where autistic traits are more affected by environmental variation. By studying these aetiological interactions in a systematic way, rather than constraining ourselves to a specific measured environment, we can develop novel hypotheses about currently unknown environmental influences.

Our findings complement previous research on geographical prevalence differences in ASD (Bakian et al., 2015; Campbell et al., 2011; Chen et al., 2008; Hoffman et al., 2017; Mazumdar et al., 2013). Similarly, alongside aetiological differences, we observe geographical variation in mean autistic trait scores. These mean differences may be linked to aetiological differences. For example, areas of greater prevalence could represent regions where the environment triggers genetic predisposition to autistic traits. This provides a basis for future research into specific geographically distributed environments that draw out or mitigate genetic or environmental risk, which could in turn be useful for population health measures seeking to reduce the impact of ASD.

From our results we can hypothesise about what these factors could be. For example, we find that there is generally higher heritability in more densely populated areas and lower heritability in less populated regions. This may suggest that urban environments draw out genetic differences in predisposition to autistic traits between people of the same ancestral background. These geographically distributed environments might include psychosocial factors such as the stress of urban living or income inequality, or aspects of the physical environment such as air pollution. This explanation fits with neuroscience literature that suggests living in an urban environment is associated with specific neural correlates in response to stress (Lederbogen et al., 2011). The literature on prevalence suggests that other potentially important geographical factors may include healthcare accessibility, diagnostic bias, parental awareness, socio-economic status, neighbourhood deprivation, area infrastructure, or green space accessibility. However, factors such as rater effects or access to healthcare are not likely to play a role in this aetiological variation as we used data from structured interviews in population representative samples, and environmental influences on prevalence are not necessarily the same as environmental influences on aetiology. For non-shared environmental influences urban-rural differences are confined to areas in and around Stockholm, the Swedish capital. Therefore, it may be that there are environments related specifically to living in or around the capital that result in decreased non-shared environmental influences compared to other areas in Sweden. As in Sweden, non-shared environmental influence in the UK shows a more complex pattern than genetic influence.

Whilst we see similarities in patterns of aetiology between Sweden and the UK for autistic traits, there are also substantial differences. There are several possible reasons

for this. It could be due to differences in the measurement of autistic traits in the cohorts, or environmental differences between the two countries. For example, differences in the level of awareness of ASD and therefore possible differential reporting in autistic traits, or differences in the physical or social environments, which may vary between countries in the same way as they do within each country. It will be important to investigate this in other countries to explore these international similarities and differences further.

When interpreting these results there are a few important points to consider. First, in some areas the effective sample size is lower than others, for example in densely populated areas the proximity of some twin pairs relative to others can weight their influence relatively highly. However, across all areas we have effective sample sizes in the thousands for both identical and fraternal twin pairs, so estimates remain reasonably precise. Second, due to how the weighting in the analyses works, i.e. participants contribute more to analysis the closer they are to the target location, this results in smoothing over the estimates for A, C and E. The amount by which results over the area are smoothed depends on the tuning parameter used in the weighting. There is a trade-off when selecting the tuning parameter between smoothing over noise and detecting real variation, or between accurately estimating variance components and accurately localising them. Here, we have chosen the tuning parameter to result in some smoothing towards the population mean, but this may mean that some larger localised variation remains undetected. In interpreting the maps, it is important to take into account both the pattern of results shown on the map, and the range of estimates shown by the histogram, while bearing in mind that the effect sizes are smoothed towards the population mean. Third, in common with the previous literature, we find

that an ADE model is often a slightly better fit to the data, but here we have fitted ACE models because the high correlation between A and D brings noise to spatial analysis due to switching between the two across locations (Rietveld, Posthuma, Dolan, & Boomsma, 2003). Instead, we interpret A here as a broad genetic component, without the usual connotation of additivity. Fourth, as with any statistical analysis, it is important to consider the assumptions of the model. For twin modelling, these include random mating within the population, that MZ and DZ twins share their environments to the same extent (at least where those environments are not genetically influenced), and that twins are representative of the general population for the traits studied (Rijsdijk & Sham, 2002). These assumptions have generally been found to be reasonable (Evans & Martin, 2000), although there is some evidence to suggest that there is assortative mating for ASD (Nordsletten et al., 2016). This would have the effect of inflating the shared environmental influences, which we find to be approximately zero across locations. For our geographical analyses we do not assume that there is no gene-environment interaction or correlation, because we are explicitly modelling them as our main point of interest.

Conclusion

Our systematic analysis shows geographical variation in genetic and non-shared environmental influences for autistic traits in both Sweden and the UK. These results will inform further studies of measured geographically distributed environments, beyond those already identified as influencing prevalence in the literature. For example, by correlating the spatial distribution of these environments with the spatial distribution of the aetiological estimates or by using formal continuous moderator models. Identifying these environments and understanding how they draw out or mask

genetic predisposition may lead to population health and social policy innovation to support people with ASD.

Acknowledgements

We gratefully acknowledge the ongoing contribution of the participants in the TEDS and CATSS and their research teams.

Funding: This work was supported in part by the UK Medical Research Council Integrative Epidemiology Unit at the University of Bristol (Grant ref: MC_UU_12013/1). ZR is supported by a Wellcome Trust PhD studentship (Grant ref: 109104/Z/15/Z). OSPD and CMAH are funded by the Alan Turing Institute under the EPSRC grant EP/N510129/1. This study was also supported by the National Institute for Health Research (NIHR) Biomedical Research Centre at the University Hospitals Bristol NHS Foundation Trust and the University of Bristol (BRC-1215-2011). The CATSS study is supported by the Swedish Foundation for International Cooperation in Research and Higher Education (STINT), the Swedish Council for Health, Working Life and Welfare, the Söderström-Königska Foundation, and the Swedish Research Council (Medicine and SIMSAM). TEDS is funded by Medical Research Council grant MR/M021475/1 to Robert Plomin. AbR is supported by the Beatrice and Samuel L Seaver Foundation.

Angelica Ronald has received a consultancy fee for writing for the National Childbirth Trust and receives an annual honorarium as joint editor of the Journal of Child Psychology and Psychiatry. H. Larsson has served as a speaker for Evolan Pharma and Shire and has received research grants from Shire; all outside the submitted work.

References

Anckarsäter, H., Larson, T., Hansson, S. L., Carlström, E., Ståhlberg, O., Gillberg, C., Råstam, M., et al. (2008). Child neurodevelopmental and behavioural problems are intercorrelated and dimensionally distributed in the general population. *The Open Psychiatry Journal*.

Anckarsäter, H., Lundström, S., Kollberg, L., Kerekes, N., Palm, C., Carlström, E., Långström, N., et al. (2011). The Child and Adolescent Twin Study in Sweden (CATSS). *Twin Research and Human Genetics*, *14*(6), 495–508.

Bai, D., Yip, B. H. K., Windham, G. C., Sourander, A., Francis, R., Yoffe, R., Glasson, E., et al. (2019). Association of Genetic and Environmental Factors with Autism in a 5-Country Cohort. *JAMA Psychiatry*, *76*(10), 1035–1043.

Bakian, A. V., Bilder, D. A., Coon, H., & McMahon, W. M. (2015). Spatial Relative Risk Patterns of Autism Spectrum Disorders in Utah. *Journal of Autism and Developmental Disorders*, *45*(4), 988–1000.

Campbell, M., Reynolds, L., Cunningham, J., Minnis, H., & Gillberg, C. (2011). Autism in Glasgow: Cumulative incidence and the effects of referral age, deprivation and geographical location. *Child: Care, Health and Development*, *39*(5), 688–694.

Chen, C. Y., Liu, C. Y., Su, W. C., Huang, S. L., & Lin, K. M. (2008). Urbanicity-related variation in help-seeking and services utilization among preschool-age children with autism in Taiwan. *Journal of Autism and Developmental Disorders*, *38*(3), 489–497.

Colvert, E., Tick, B., McEwen, F., Stewart, C., Curran, S. R., Woodhouse, E., Gillan, N., et al. (2015). Heritability of autism spectrum disorder in a UK population-based twin sample. *JAMA Psychiatry*, *72*(5), 415–423.

Constantino, J. N., & Todd, R. D. (2003). Autistic Traits in the General Population. *Archives of General Psychiatry*, *60*(5), 524.

Davis, O. S. P., Haworth, C. M. A., Lewis, C. M., & Plomin, R. (2012). Visual analysis of geocoded twin data puts nature and nurture on the map. *Molecular Psychiatry*, *17*(9), 867–874.

Evans, D. M., & Martin, N. G. (2000). The validity of twin studies. *Genescreen*, *1*(2), 77.

Hansson, S. L., Röjvall, A. S., Rastam, M., Gillberg, C., Gillberg, C., & Anckarsäter, H. (2005). Psychiatric telephone interview with parents for screening of childhood autism - Tics, attention-deficit hyperactivity disorder and other comorbidities (A-TAC): Preliminary reliability and validity. *British Journal of Psychiatry*, *187*(SEPT.), 262–267.

Haworth, C. M. A., Davis, O. S. P., & Plomin, R. (2013). Twins Early Development Study (TEDS): A Genetically Sensitive Investigation of Cognitive and Behavioural Development From Childhood to Young Adulthood. *Twin Research and Human Genetics*, *16*(1), 117–125.

Hoffman, K., Weisskopf, M. G., Roberts, A. L., Raz, R., Hart, J. E., Lyall, K., Hoffman, E. M., et al. (2017). Geographic Patterns of Autism Spectrum Disorder Among Children of Participants in Nurses ' Health Study II. *American Journal of Epidemiology*, *186*(7),

834–842.

Idring, S., Lundberg, M., Sturm, H., Dalman, C., Gumpert, C., Rai, D., Lee, B. K., et al. (2015). Changes in Prevalence of Autism Spectrum Disorders in 2001–2011: Findings from the Stockholm Youth Cohort. *Journal of Autism and Developmental Disorders*, *45*(6), 1766–1773.

Larson, T., Anckarsäter, H., Gillberg, C., Ståhlberg, O., Carlström, E., Kadesjö, B., Råstam, M., et al. (2010). The Autism-Tics, AD/HD and other Comorbidities inventory (A-TAC): further validation of a telephone interview for epidemiological research. *BMC psychiatry*, *10*(1), 1.

Larson, T., Lundström, S., Nilsson, T., Selinus, E. N., Råstam, M., Lichtenstein, P., Gumpert, C. H., et al. (2013). Predictive properties of the A-TAC inventory when screening for childhood-onset neurodevelopmental problems in a population-based sample. *BMC Psychiatry*, *13*, 233.

Lauritsen, M. B., Astrup, A., Pedersen, C. B., Obel, C., Schendel, D. E., Schieve, L., Yeargin-Allsopp, M., et al. (2014). Urbanicity and autism spectrum disorders. *Journal of Autism and Developmental Disorders*, *44*(2), 394–404.

Lederbogen, F., Kirsch, P., Haddad, L., Streit, F., Tost, H., Schuch, P., Wüst, S., et al. (2011). City living and urban upbringing affect neural social stress processing in humans. *Nature*, *474*(7352), 498–501.

Lundström, S., Chang, Z., Råstam, M., Gillberg, C., Larsson, H., Anckarsäter, H., & Lichtenstein, P. (2012). Autism spectrum disorders and autisticlike traits: Similar

etiology in the extreme end and the normal variation. *Archives of General Psychiatry*, *69*(1), 46–52.

Lundström, S., Reichenberg, A., Anckarsäter, H., Lichtenstein, P., & Gillberg, C. (2015). Autism phenotype versus registered diagnosis in Swedish children: Prevalence trends over 10 years in general population samples. *BMJ*, *350*, h1961.

Lyall, K., Croen, L., Daniels, J., Fallin, M. D., Ladd-Acosta, C., Lee, B. K., Park, B. Y., et al. (2017). The Changing Epidemiology of Autism Spectrum Disorders. *Annual Review of Public Health*, *38*(1), 81–102.

Magnusson, P. K. E., Almqvist, C., Rahman, I., Ganna, A., Viktorin, A., Walum, H., Halldner, L., et al. (2012). The Swedish Twin Registry: Establishment of a Biobank and Other Recent Developments. *Twin Research and Human Genetics*, *16*(01), 317–329.

Mårland, C., Lichtenstein, P., Degl'Innocenti, A., Larson, T., Råstam, M., Anckarsäter, H., Gillberg, C., et al. (2017). The Autism-Tics, ADHD and other Comorbidities inventory (A-TAC): Previous and predictive validity. *BMC Psychiatry*, *17*(1), 403.

Mazumdar, S., Winter, A., Liu, K. Y., & Bearman, P. (2013). Spatial clusters of autism births and diagnoses point to contextual drivers ofincreased prevalence. *Social Science and Medicine*, *95*, 87–96.

Nordsletten, A. E., Larsson, H., Crowley, J. J., Almqvist, C., Lichtenstein, P., & Mataix-Cols, D. (2016). Patterns of nonrandom mating within and across 11 major psychiatric disorders. *JAMA Psychiatry*, *73*(4), 354–361.

Rietveld, M. J. H., Posthuma, D., Dolan, C. V., & Boomsma, D. I. (2003). ADHD: Sibling interaction or dominance: An evaluation of statistical power. *Behavior Genetics*, *33*(3), 247–255.

Rijsdijk, F. V, & Sham, P. C. (2002). Analytic approaches to twin data using structural equation models. *Briefings in bioinformatics*, *3*(2), 119–133.

Robinson, E. B., Koenen, K. C., McCormick, M. C., Munir, K., Hallett, V., Happé, F., Plomin, R., et al. (2011). Evidence that autistic traits show the same etiology in the general population and at the quantitative extremes (5%, 2.5%, and 1%). *Archives of General Psychiatry*, *68*(11), 1113–1121.

Robinson, E. B., St Pourcain, B., Anttila, V., Kosmicki, J. A., Bulik-Sullivan, B., Grove, J., Maller, J., et al. (2016). Genetic risk for autism spectrum disorders and neuropsychiatric variation in the general population. *Nature Genetics*, *48*(5), 552–555.

Ronald, A, Larsson, H., Anckarsäter, H., & Lichtenstein, P. (2011). A twin study of autism symptoms in Sweden. *Molecular Psychiatry*, *16*(10), 1039–1047.

Ronald, Angelica, Happé, F., Bolton, P., Butcher, L. M., Price, T. S., Wheelwright, S., Baron-Cohen, S., et al. (2006). Genetic heterogeneity between the three components of the autism spectrum: A twin study. *Journal of the American Academy of Child and Adolescent Psychiatry*, *45*(6), 691–699.

Sandin, S., Lichtenstein, P., Kuja-Halkola, R., Hultman, C., Larsson, H., & Reichenberg, A. (2017). The Heritability of Autism Spectrum Disorder. *JAMA*, *318*(12), 1182–1184.

Scott, F. J., Baron-Cohen, S., Bolton, P., & Brayne, C. (2002). The CAST (Childhood Asperger Syndrome Test): preliminary development of a UK screen for mainstream primary-school-age children. *Autism*, *6*(1), 9–31.

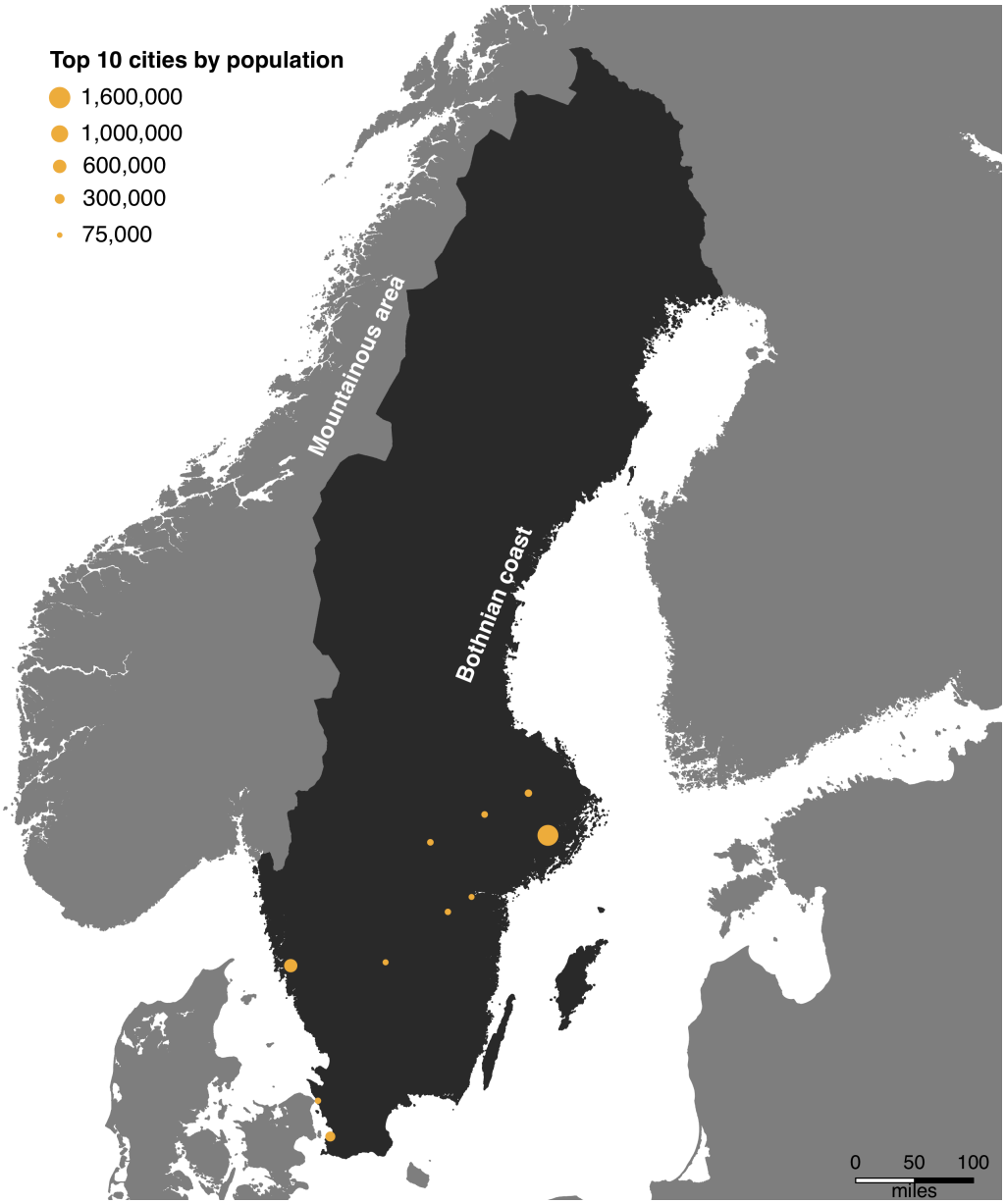
Seltzer, M. M., Shattuck, P., Abbeduto, L., & Greenberg, J. S. (2004). Trajectory of development in adolescents and adults with autism. *Mental Retardation and Developmental Disabilities Research Reviews*, *10*(4), 234–247.

Tick, B., Bolton, P., Happé, F., Rutter, M., & Rijsdijk, F. (2016). Heritability of autism spectrum disorders: A meta-analysis of twin studies. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *57*(5), 585–595.

Vassos, E., Agerbo, E., Mors, O., & Bøcker Pedersen, C. (2016). Urban-rural differences in incidence rates of psychiatric disorders in Denmark. *British Journal of Psychiatry*, *208*(5), 435–440.

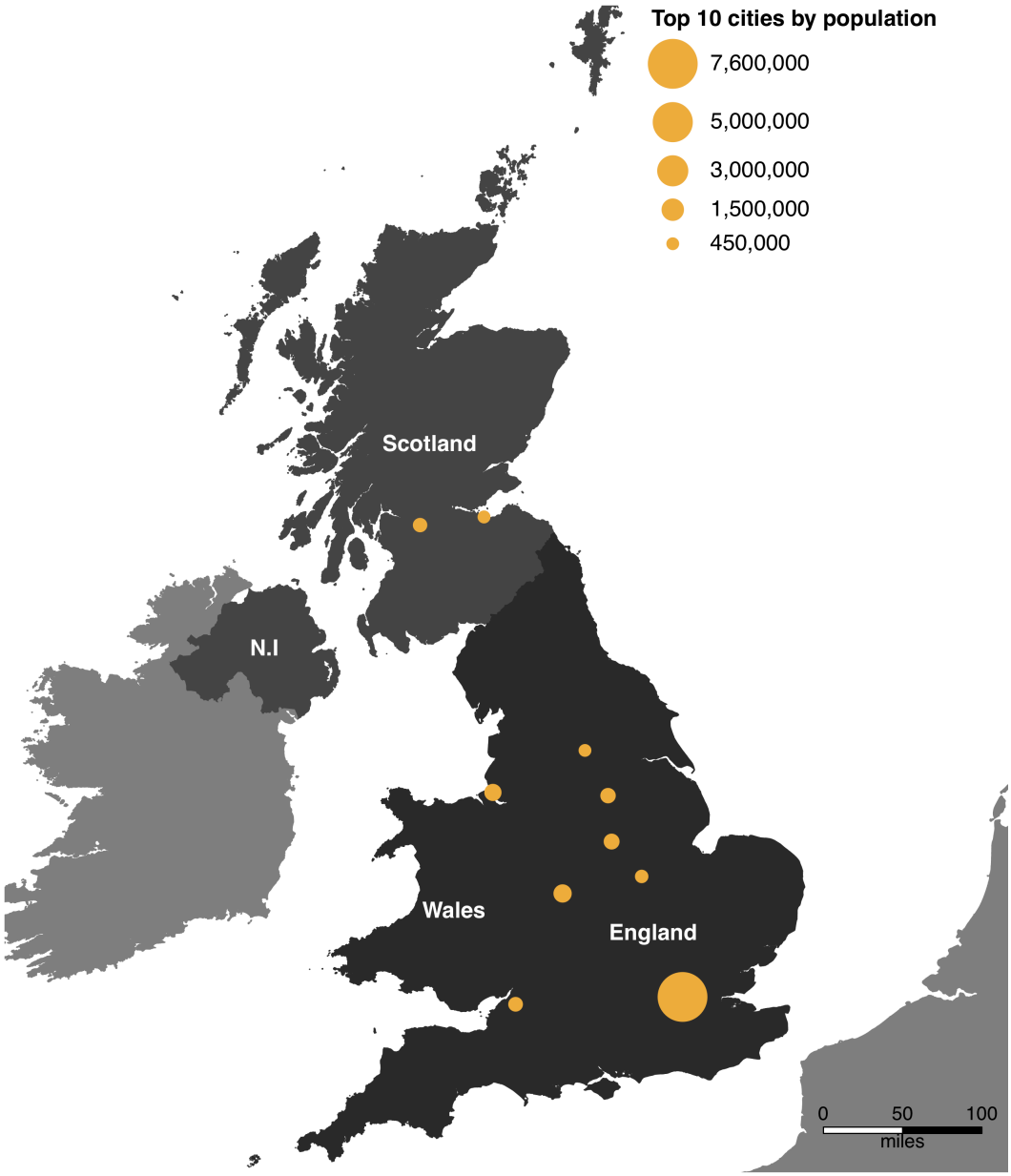
Wu, J., & Jackson, L. (2017). Inverse relationship between urban green space and childhood autism in California elementary school districts. *Environment International*, *107*, 140–146.

Figure 1. Map of Sweden with top 10 cities by population



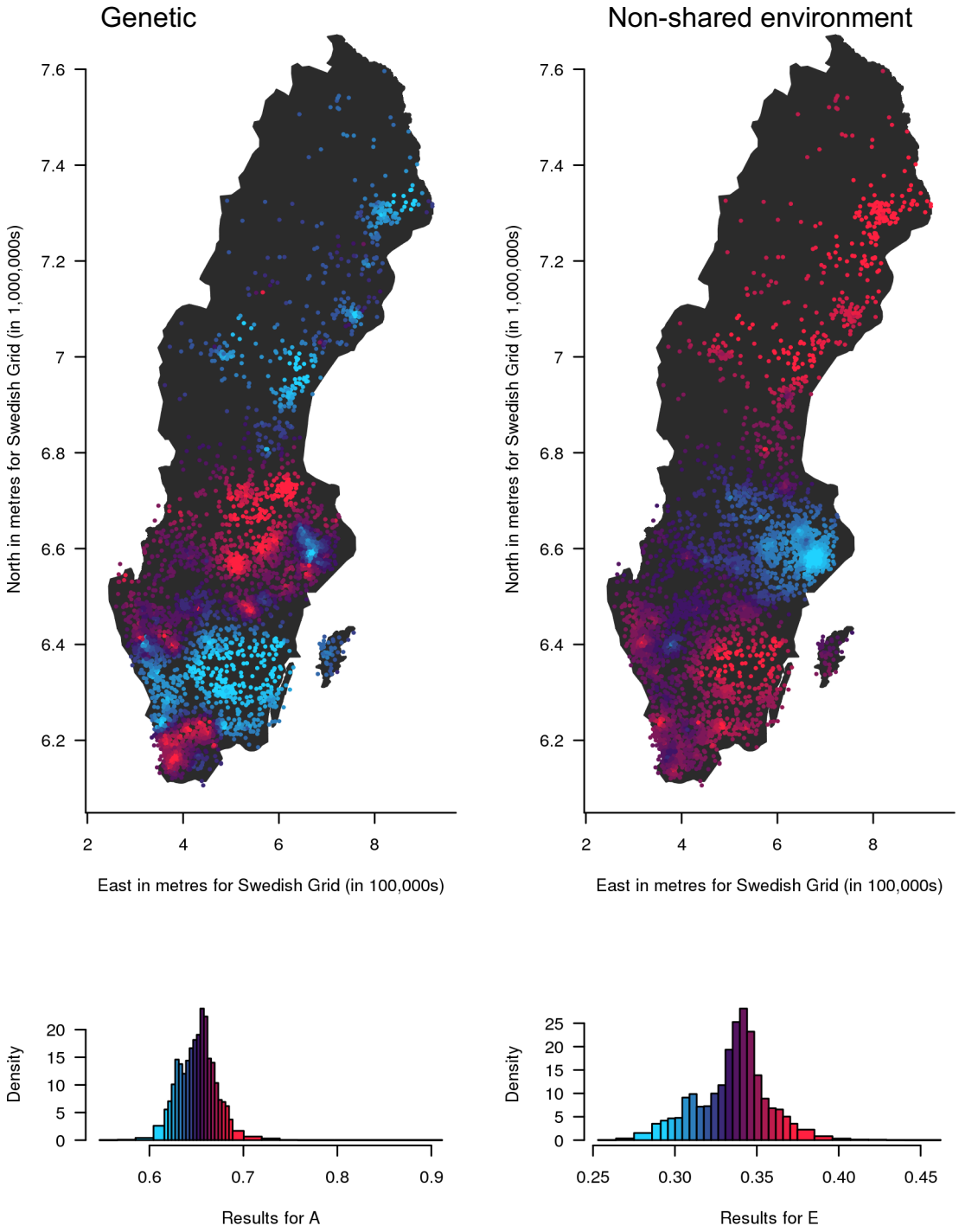
Sweden is shown in dark grey with the surrounding countries in a lighter grey. The top 10 most populated cities are indicated by orange circles, where the area reflects the population.

Figure 2. Map of the UK with top 10 cities by population



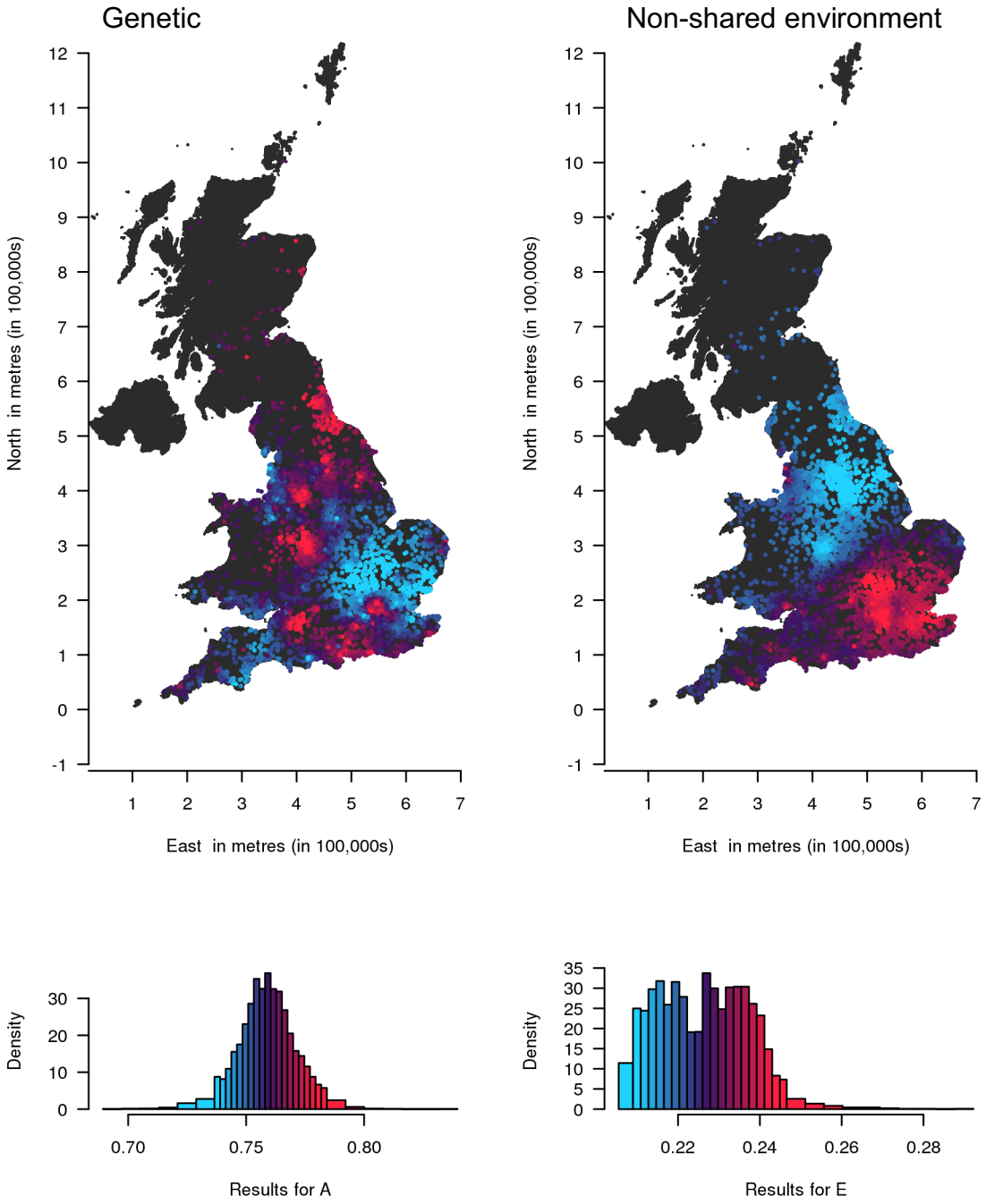
The recruitment area (England and Wales) for the Twins Early Development Study (TEDS) is shown in dark grey, with the rest of the UK (Scotland and Northern Ireland [N.I]) in a lighter grey. Other countries are shown in the lightest shade of grey. The top 10 most populated cities are indicated by orange circles, where the area reflects the population.

Figure 3. Mapping genetic (A) and non-shared environmental (E) influences on autistic traits in Sweden suggests that genetic variation is more influential in more densely populated areas



Geographical variation in genetic (A) and non-shared (E) influences on childhood autistic traits in Sweden (results overlaid on an outline of the SAMS areas). The contributions of A and E range from low (blue) to high (red). The histograms show the distribution of the estimates, coloured in the same way as the points on the map. The estimates are not standardised and are therefore not constrained to add up to one. A higher-resolution interactive version of this map is available at https://dynamicgenetics.github.io/spACEjs/.

Figure 4. Genetic influences for autistic traits in the UK appear greater in more densely populated areas, although patterns of non-shared environmental influences follow a north-south divide



Geographical variation in genetic (A) and non-shared (E) influences on childhood autistic traits in the UK. The contributions of A and E range from low (blue) to high (red). The histograms show the distribution of

the estimates, coloured in the same way as the points on the map. The A and E estimates are not standardised and are therefore not constrained to add up to one at each location. A higher-resolution interactive version of this map is available at https://dynamicgenetics.github.io/spACEjs/.