Arc requires PSD95 for assembly into postsynaptic complexes involved with brain disease and intelligence

~~Proteomic and genetic dissection of Arc complexes shows a convergent mechanism for mental disorders and intelligence.~~

Formatted: Font: 14 pt

Formatted: Font: 18 pt

Esperanza Fernández[1,12,14], Mark O. Collins[2,13], René A.W. Frank[1,3], F. Zhu[1,4], Maksym V. Kopanitsa[1,5], Jess Nithianantharajah[1,4], Sarah A. Lemprière[4], David Fricker[5], Kathryn A. Elsegood[1,4], Catherine L. McLaughlin[4], Mike D.R. Croning[4], Colin Mclean[6], J. Douglas Armstrong[6], W. David Hill[7], Ian J. Deary[7], Giulia Cencelli[12], Claudia Bagni[12], Menachem Fromer[8,9,10], Shaun M. Purcell[10], Andrew Pocklington[11], Jyoti S. Choudhary[2], Noboru H. Komiyama[5], Seth G. N. Grant[1,4]

Affiliations:

[1] Genes to Cognition Programme, The Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, United Kingdom.
[2] Proteomic Mass Spectrometry, The Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, United Kingdom.
[3] Medical Research Council Laboratory of Molecular Biology, Cambridge, United Kingdom.
[4] Genes to Cognition Programme, Centre for Clinical Brain Science, University of Edinburgh, Edinburgh, United Kingdom.
[5] Synome Ltd., Moneta Building, Babraham Research Campus, Cambridge, United Kingdom.
[6] School of Informatics, Institute for Adaptive and Neural Computation, University of Edinburgh, United Kingdom.
[7] Centre for Cognitive Ageing and Cognitive Epidemiology, Department of Psychology, University of Edinburgh, United Kingdom.
[8] Analytic and Translational Genetics Unit, Massachusetts General Hospital, Boston, Massachusetts 02114, USA
[9] Stanley Center for Psychiatric Research, Broad Institute of MIT and Harvard, Cambridge, Massachusetts, 02142, USA.
[10] Division of Psychiatric Genomics, Department of Psychiatry, Icahn School of Medicine at Mount Sinai, New York, New York 10029, USA.
[11] Institute of Psychological Medicine & Clinical Neurosciences, University of Cardiff, Cardiff, Wales, United Kingdom.
[12] KU Leuven, Center for Human Genetics and Leuven Institute for Neurodegenerative Diseases (LIND), and VIB Center for the Biology of Disease, Leuven, Belgium
[13] Department of Biomedical Science, The Centre for Membrane Interactions and Dynamics, University of Sheffield, Sheffield, United Kingdom.
[14] Current address: VIB-UGent Center for Medical Biotechnology, Gent, Belgium

Corresponding author:

Seth G.N. Grant
e-mail: Seth.Grant@ed.ac.uk
Ph: +44 (0)131 242 7984

SUMMARY

Arc is an activity regulated neuronal protein yet little is known about its protein interactions, assembly into multiprotein complexes, role in human disease and cognition. We applied an integrated proteomic and genetic strategy using targeted tagging of a Tandem Affinity Purification (TAP) tag and Venus fluorescent protein into the endogenous Arc gene in mice, biochemical and proteomic characterization of native complexes in wild type and knockout mice, and human genetic analyses of disease and intelligence. TAP tagging enabled efficient purification of complexes and identification of many novel Arc-interacting proteins, of which PSD95 was the most abundant. PSD95 was essential for Arc assembly into 1.5 MDa complexes and activity-dependent recruitment to excitatory synapses. Integrating human genetic data with proteomic data showed postsynaptic Arc-PSD95 complexes are enriched in schizophrenia, intellectual disability, autism and epilepsy mutations and normal variants in intelligence. Arc-PSD95 postsynaptic complexes are a target for normal and pathological genetic variants impacting on human cognitive function.~~Arc-PSD95 postsynaptic complexes are a molecular substrate for the convergence of normal and pathological genetic variants impacting on human cognitive function.~~

INTRODUCTION

Arc/Arg3.1 was originally identified as a cytoskeletal protein encoded by an mRNA that was rapidly transcribed following synaptic activity and transported to dendrites (Link et al., 1995; Lyford et al., 1995; Moga et al., 2004; Steward et al., 1998). Many forms of neuronal activation induce Arc: synaptic stimulation including long-term potentiation (Guzowski et al., 2000), metabotropic glutamate receptor-dependent long-term depression, (Jakkamsetti et al., 2013; Park et al., 2008; Waung et al., 2008), homeostatic scaling of AMPA receptors (Gao et al., 2010; Korb et al., 2013; Okuno et al., 2012; Shepherd et al., 2006), generalized neuronal activity induced by seizures (Link et al., 1995) as well as various behavioral stimuli (memory- and experience-related behavioral patterns) (Daberkow et al., 2007; Gao et al., 2010; Guzowski et al., 1999; Jakkamsetti et al., 2013; Kelly and Deadwyler, 2003; Miyashita et al., 2009; Vazdarjanova and Guzowski, 2004; Vazdarjanova et al., 2006; Wibrand et al., 2012), and visual stimuli (Wang et al., 2006). Knockout or knockdown of Arc results in impaired synaptic plasticity and hippocampus-dependent learning and behavior phenotypes reminiscent of schizophrenia (Guzowski et al., 2000; Managò et al., 2016; McCurry et al., 2010; Plath et al., 2006; Wang et al., 2006).

Arc is mainly localized at postsynaptic sites of excitatory synapses (Moga et al., 2004). The proteome of the postsynaptic terminal of excitatory synapses of vertebrate species contains a highly conserved set of ~1000 protein types (Bayes et al., 2012; Bayes et al., 2017; Bayes et al., 2011b; Distler et al., 2014) organized into >200 multiprotein complexes (Frank and Grant, 2017; Frank et al., 2016; Frank et al., 2017b). The multiprotein

complexes are organized into a hierarchy of complexes and supercomplexes (complexes of complexes) and the prototype supercomplex is formed by PSD95 (Fernandez et al., 2009; Frank et al., 2016; Frank et al., 2017b; Husi and Grant, 2001; Husi et al., 2000). Arc was found to be associated with PSD95 (Fernandez et al., 2009; Frank et al., 2016; Frank et al., 2017a; Husi et al., 2000) and genetic studies show that absence of either PSD95 or Arc leads to enhanced Long-Term Potentiation (LTP) and impaired hippocampus-dependent learning (Migaud et al., 1998; Plath et al., 2006). Biochemical purification and mouse genetic experiments show dimers of PSD95 assemble with multiple complexes including NMDA receptors, potassium channels, signaling and adhesion proteins. These are not all found within a single supercomplex, but are within an extended family of PSD95 supercomplexes ranging in size from 1–3 MDa (Frank and Grant, 2017; Frank et al., 2016; Frank et al., 2017b). A large-scale mouse genetic screen of >50 postsynaptic proteins found that PSD95 and its close interacting proteins had the strongest phenotypes in synaptic electrophysiology and behavior, indicating that PSD95 supercomplexes are crucial components of the postsynaptic terminal of excitatory synapses (Komiyama et al., submitted; Kopanitsa et al., submitted). Arc has also been proposed to interact with the endocytic machinery (Dynamin, Endophilin-2 and -3) (Chowdhury et al., 2006; Rial Verde et al., 2006; Shepherd et al., 2006). However, Arc multiprotein complexes have not been purified and systematically studied using proteomic mass spectrometry and thus the identify of its interacting partners and composition of Arc complexes remains poorly understood.

Characterizing protein complexes in synapses is technically challenging and the use of gene-tagging of endogenous proteins in the mouse has greatly facilitated purification of intact native complexes and visualization of their subcellular localization, and has many advantages over *in vitro* and recombinant methods (Broadhead et al., 2016; Fernandez et al., 2009; Frank and Grant, 2017; Frank et al., 2016). The effect of mutations on complexes and neuronal activation can be combined in mice carrying knockin gene tags, and proteins that are predicted to be largely unstructured and form multivalent interactions, such as Arc (Xue et al., 2010), can be studied in their native context. These advantages have been illustrated by the purification of native NMDA receptor and PSD95 complexes, where a Tandem-Affinity Purification (TAP) tag was inserted into the N-terminus of the GluN1 subunit and C-terminus of PSD95 by genome engineering (Fernandez et al., 2009; Frank et al., 2016). Purification revealed NMDA receptors and PSD95 were in ~1.5 MDa supercomplexes with channel subunits, PSD95 and PSD93 as major components. Genetic dissection *in vivo* using mutant mice showed an essential tripartite requirement for PSD95, PSD93 and the GluN2B cytoplasmic domain (Frank et al., 2016). This tripartite interaction was not previously detected using *in vitro* methods which typically rely on binary protein interactions. Moreover, like Arc, the GluN2B cytoplasmic domain is predicted to be a structurally unfolded/disordered domain (Ryan et al., 2008), and these domains lack stable tertiary structure, and undergo disorder-to-order transitions upon binding or changes in phosphorylation (Bah et al., 2015; Gibbs et al., 2017). We therefore considered that Arc was well suited to the strategy of gene tagging and genetic dissection.

Genetics has been a powerful approach for studying the function of multiprotein complexes in many prokaryotic, eukaryotic and metazoan organisms, including humans, where disease-causing mutations have been mapped to protein complexes (Babu et al., 2014; Lu et al., 2013; Vidal et al., 2011). Moreover, in recent years, a large number of mutations that disrupt postsynaptic proteins in humans have been identified and found to cause many psychiatric, neurological and developmental disorders (Bayes et al., 2014; Bayes et al., 2011a; Brose et al., 2010; Fromer et al., 2014; Grant, 2012; Grant et al., 2005; Grant, 2013; Kirov et al., 2012; Pocklington et al., 2006; Purcell et al., 2014). Although mutations in the human *ARC* gene have not been directly linked to any mental disorder, using preliminary proteomic data on Arc interacting proteins, the proteins in Arc complexes were found to be enriched in disruptive mutations (Purcell et al., 2014), *de novo* copy-number variants (CNVs) (Kirov et al., 2012), non-synonymous *de novo* single nucleotide variants (SNVs) and small insertions and deletions (indels) (Fromer et al., 2014) in schizophrenia cases. ARC protein has been described to accumulate at synapses in Angelman syndrome (Greer et al., 2010) and increased and/or decreased in several animal models of Alzheimer's disease and patient-derived cells (for review see (Kerrigan and Randall, 2013). These data suggest that Arc is a component of protein complexes that are involved with human cognitive disorders.

In this paper, we have conducted an extensive proteomic and genetic dissection of Arc protein complexes, which is a generic strategy suitable for the characterization of potentially any synaptic protein. We have focused on the following four challenges: i) the isolation of native multiprotein complexes from brain tissue, ii) the visualization of the

endogenous protein using genetic tagging, iii) genetic dissection of protein complex organization and localization using mouse genetic models, iv) genetic dissection of complexes using human genetic data including human disease and cognitive phenotypes. Here we demonstrate that this integrated proteomic and genetic strategy reveals ~~important new~~ insights into Arc's role in biology and the synaptic basis of mental disorders and intelligence.

RESULTS

TAP tagging and proteomic analysis of endogenous Arc complexes

To label and isolate endogenous Arc protein we engineered knockin mice (Arc^{TAP}) harboring a Tandem Affinity Purification (TAP) tag fused to Arc's carboxyl terminus (Figure 1A; Figures S1, S2). Mice carrying the TAP tag showed no detectable alterations in levels or localization of Arc in the brain or in hippocampal synaptic physiology (Figure S3 and S4). Native Arc complexes were detected by immunoblotting of brain extracts separated on blue-native PAGE (BNP), which showed a major band of median mass ~1.5 MDa with several additional minor species ranging from ~200-700 kDa (Figure 1B).

The TAP tag was used to isolate Arc complexes directly from mouse forebrain tissue using a highly efficient purification protocol (recovering >70% Arc) (Figure 1C,D) and their composition determined using LC-MS/MS. The single-step purification yielded 102 high-confidence proteins whereas the more stringent tandem-purification protocol recovered a subset of 39 proteins (34/39 were uncovered by the single-step purification)(Experimental procedures, Table S1) (http://www.genes2cognition.org/publications/tap-arc). Eight out of

14 previously reported Arc interactors were found among the 107 high-confidence proteins indicating that 99 were novel interactors (Table S2). Among the 107 high-confidence proteins, 72 proteins contain the Arc-N lobe consensus motif P[STVILMKR][FYH] (Zhang et al., 2015) revealing a strong network of direct interactors (Table S1). Comparison of mouse and human show 87% (92/107) of Arc interacting proteins were conserved between species (Table S3), 70% (1012 in human of 1447 in mice) of protein-protein interactions were conserved (Table S4) and the Arc interactome was enriched (72%) in proteins in the human postsynaptic complexes found in Bayes et al., (2012) (Table S4). Together these results suggest we have defined a robust Arc complex and interactome that is highly conserved between mouse and human.

PSD95 was the most abundant Arc interacting protein: using iBAQ quantification (Schwanhausser et al., 2011) of the single-step purification it showed ~1:1 stoichiometry with Arc (Figure 1E, Table S5) and in the tandem-purification it represented 57% of the Arc interactome (Table S6). Reciprocal immunoprecipitations show PSD95 assembles into Arc complexes from early developmental stages (P11) in hippocampus and cortex (Figure S5). The Dlg gene family of adaptor/scaffold proteins, comprising four paralogs (SAP97/*Dlg1*, PSD93/*Dlg2*, SAP102/*Dlg3*, PSD95/*Dlg4*), was the most abundant of eleven protein classes recovered, suggesting they play a principal role in regulating Arc function (Figure S6; Tables S7 and S8). Specificity of interaction between Arc and specific Dlg paralogs was suggested by the finding that PSD93 and SAP97 were also highly abundant, whereas SAP102 was not detected in the Arc interactome (confirmed using reciprocal immunoprecipitation, Figure S7). Forty nine percent of Arc interacting proteins

were known PSD95 interactors and particularly enriched in membrane proteins including NMDA and AMPA receptors (Table S1).

Consistent with their coassembly with Arc, the NMDA receptor, PSD95 and PSD93 were also shown to reside in 1.5 MDa supercomplexes (Frank et al., 2016). To compare the relative abundance of Arc and PSD95 in 1.5MDa supercomplexes we immunoblotted $Arc^{TAP/+}$ and $PSD95^{TAP/+}$ brain extracts separated by BNP with Flag antibodies (Figure 1F). PSD95 was ~20 fold more abundant than Arc indicating ~5% of PSD95 complexes contain Arc. Pull down of PSD95 complexes using the TAP tag recovered 65% of PSD95 and depleted 37% of Arc, indicating that ~58% of Arc is in PSD95 supercomplexes (data not shown). Together, these data indicate that PSD95 is the major interacting protein of Arc and that a subset of the postsynaptic 1.5 MDa PSD95 supercomplexes contain Arc.

Arc postsynaptic localization requires PSD95

How Arc is localized to the postsynaptic terminal is unknown. To address this question we asked if members of the Dlg scaffold proteins were required *in vivo*, using mice carrying knockout mutations. In hippocampal extracts from PSD95, PSD93 and SAP102 knockout mice (SAP97 knockout mice are unviable) we found that Arc protein levels were reduced (35.0 ± 17.1% of WT; $p < 0.01$) in PSD95 knockout mice but not in either PSD93 or SAP102 knockout mice (Figure 2A and Figure S8). A dramatic loss of dendritic staining of TAP tagged Arc was observed in hippocampal sections from PSD95 knockout mice (Figure 2B). Furthermore, synaptosomes from PSD95 knockout mice also showed a major reduction in Arc (Figure 2C). We also examined BNP immunoblots from PSD95

knockouts and found that 1.5 MDa Arc complexes were severely diminished, with a weak residual signal after long exposure of the gel (Figure 2D). Thus, PSD95 is specifically required to localize Arc to the postsynaptic terminal.

To visualize endogenous Arc protein we created Arc^{Venus} knockin mice, using a similar design strategy to the Arc^{TAP} mice, where the Venus fluorescent protein was fused to the C-terminus of Arc (Figure S1B). Mice carrying the Venus tag showed no detectable alterations in hippocampal synaptic physiology (Figure S9). We bred Arc-Venus mice with PSD95 knockouts to generate compound transgenic mice (Arc^{Venus}x$PSD95^{-/-}$) and asked if kainic acid induced neuronal activity (Li et al., 2005) would drive Arc to the synapse and whether this required PSD95. In the absence of PSD95, Arc-Venus failed to localize to the synapse (Figure 2E). These results demonstrate that PSD95 is required for the postsynaptic localization of Arc into 1.5 MDa complexes in the steady state and following induction by neuronal activity.

Proteomic analysis of Arc complexes in mice lacking PSD95

We reasoned that by genetic removal of PSD95, we could identify those Arc interacting proteins that were most dependent on PSD95. We bred $Arc^{TAP/TAP}$ with PSD95 knockout mice ($Arc^{TAP/TAP}$/$PSD95^{-/-}$) mice and analyzed their Arc interactome using quantitative proteomic methods (Figure 3A; Figure S10 and Table S9). As shown in Figure 3A, Arc interacting proteins separated into two broad subgroups: depleted and enriched proteins (see red and green proteins respectively). Seventy percent of depleted proteins were PSD95 interacting proteins, including PSD93. As shown by immunoblots of BNPs, PSD93

remained in 1.5 MDa complexes in PSD95$^{-/-}$ mice (Figure 2D and (Frank et al., 2016). Absence of PSD93 did not affect the interaction of Arc with PSD95 (Figure S11). The most significant GO Biological Process (BP) terms in the depleted proteome are synaptic transmission (p-value=1.23x10-11), cell-cell signaling (p-value=3.94x10-9), and modulation of synaptic transmission (p-value=1.35x10-7) highlighting the functional importance of the depleted proteins (Table S8).

Amongst the 12 most enriched proteins in the Arc complexes isolated from PSD95 mutant mice were SAP97 and structural proteins including those with a potential role in cell growth and adhesion (Claudin11, Lgi1). A network graph of the interactions of the enriched and depleted proteins sets is shown in Figure 3B. The internal network consists of 26 proteins and 44 interactions (visualised using Visone, (Brandes, 2012). Taken together these proteomic and *in vivo* genetic studies show that Arc is tethered to postsynaptic 1.5 MDa signaling complexes containing PSD95 and when these complexes are abolished in PSD95 mutants, Arc is found associated with cytoskeletal and structural-related proteins. Thus, Arc is partitioned into either the PSD95 supercomplexes in the postsynaptic terminal or into cytoskeletal complexes.

Arc complexes in disease

Proteins within the postsynaptic proteome are assembled into complexes and supercomplexes (Frank et al., 2016) and this supramolecular organization is of crucial importance in human genetic disorders since it is a mechanism by which the many different gene products functionally converge. We therefore combined our proteomic datasets with human genetic datasets to understand the importance of Arc interacting proteins in human disease.

Preliminary proteomic data from Arc^{TAP} mice has previously been used to implicate the disruption of ARC complexes in human psychiatric disorders. The first such study revealed that components of Arc complexes were enriched in *de novo* CNVs from individuals with schizophrenia (Kirov et al., 2012), with subsequent studies finding enrichment for rare point mutations in individuals with schizophrenia, autism and intellectual disability (ID) (Fromer et al., 2014; Kirov et al., 2012; Purcell et al., 2014). Here we extend this analysis using complete sets of Arc interacting proteins and additional genetic datasets including epilepsy and healthy control *de novo* datasets.

We first sought to replicate the initial finding of (Kirov et al., 2012) with the comprehensive Arc interactome. Utilizing the same genetic dataset as (Kirov et al., 2012) we found that *de novo* CNVs from schizophrenia probands were enriched for Arc complex genes compared to *de novo* CNVs from unaffected individuals ($P = 0.0047$). Arc interactors whose association with Arc is depleted in PSD95 knockout mice largely drove this enrichment ($P = 0.0165$), indicating the importance of the postsynaptic 1.5 MDa

complexes. We next investigated enrichment of the Arc interactome for rare point mutations and indels contributing to brain disorders, using exome sequencing data from a case/control schizophrenia study (Purcell et al., 2014) and *de novo* studies performed in cohorts of schizophrenia, autism, intellectual disability and epilepsy (see Supplemental Experimental Procedures). Combining evidence from each of these independent datasets we found strong support for the enrichment of both nonsynonymous (NS) and loss-of-function (LoF) disease-related mutations amongst Arc interactors (P = 9.01 x 10^{-12} and 2.051 x 10^{-7} respectively, Table 1). All 5 datasets contributed to this enrichment (Table 1, Table S10 and Table S11), indicating that disruption of Arc complexes may contribute to a wide range of brain disorders. Consistent with the analysis of *de novo* CNVs, much of the enrichment in LoF and NS mutations was attributable to Arc interactors whose expression is altered in the PSD95$^{-/-}$ mouse. This suggests that it is postsynaptic Arc-PSD95 complexes and not cytoplasmic Arc complexes that are relevant to these disorders.

Arc complexes in normal variation in human intelligence

Although the role in cognition for Arc, PSD95 and their interacting proteins is well established from studies of mutations in mice (Fernandez et al., 2009; Fitzgerald et al., 2014; Husi et al., 2000; Komiyama et al., 2002; McCurry et al., 2010; Migaud et al., 1998; Nithianantharajah et al., 2013; Plath et al., 2006; Ryan et al., 2013), and mutations in humans cause cognitive disorders, and enrichment analysis of Arc interacting proteins for Mammalian Phenotype (MP) terms shows 48 enriched terms ($P < 0.01$) associated with abnormal synaptic and cognitive functions (Table S12), much less is known about

the relevance to normal variation in human cognition. We therefore asked whether common genetic variation in Arc complexes was associated with common variation in general cognitive ability (known as intelligence or *g*) using the genome-wide association study (GWAS) on intelligence from the five cohorts (n = 3511) that make up the Cognitive Ageing in England and Scotland (CAGES) consortium (Davies et al., 2011; Hill et al., 2014a). The five cohorts are the Lothian Birth Cohort of 1921 and 1936 (Deary et al., 2012), the Aberdeen Birth Cohort of 1936 (Whalley et al., 2011), and the Manchester and Newcastle Longitudinal Studies of Cognitive Ageing (Rabbitt et al., 2004), which together consist of a total of 3511 healthy middle to old aged individuals who all live independently in the community. The measure of general cognitive ability was taken from the GWAS previously conducted by (Hill et al., 2014a) (Supplemental Experimental Procedures). In order to determine if there was a greater weight of evidence for association between the Arc gene set and general cognitive ability a two-stage enrichment test was used. Firstly, single nucleotide polymorphisms (SNPs) were assigned to autosomal genes and a gene based statistic was derived (Liu et al., 2010a). Secondly, the p-values of the gene based statistics were –log(10) transformed before Gene Set Enrichment Analysis (GSEA) (Subramanian et al., 2005), a competitive test of enrichment was used Gene Set Enrichment Analysis (GSEA) (Subramanian et al., 2005). The results of the gene based analysis are shown in Table S13 where eight genes were nominally significant in CAGES and nine in BATS (the Brisbane Adolescent Twin Study) (BATS). The most significant gene in the BATS cohort (*PRRT1*, $P = 0.00732$) was also nominally significant in CAGES, $P = 0.03797$. The results of the enrichment analysis show that common genetic variation in Arc complex proteins show nominally significant association ($P = 0.0473$) with

intelligence compared with control gene sets. A replication study using the summary data of a GWAS conducted on intelligence (Hill et al., 2014a) the BATS (n = 2062) (de Zubicaray et al., 2008; Wright et al., 2001; Wright and Martin, 2004) also showed a significant enrichment (*P* = 0.0247) confirming the results found in the CAGES consortium. This significant enrichment shows that common genetic variation in the genes encoding Arc complex proteins is associated with the normal variation in human intelligence differences.

DISCUSSION

We have developed and demonstrated an integrated proteomic and genetic strategy that reveals ~~important new~~ insights into Arc's role in biology, the synaptic basis of mental disorders and intelligence. Multiple genetic and genome engineering methods were combined to isolate native Arc complexes, identify their constituents, determine the mechanism of assembly and localization to the postsynaptic terminal, and identify multiple diseases and mutations that converge on the complexes.

Arc protein is principally housed within 1.5 MDa complexes and proteomic mass spectrometry identified many novel Arc interacting proteins of which PSD95 was the most abundant. PSD95 and Arc coassemble into 1.5 MDa supercomplexes and knockout of PSD95 abolishes these complexes and severely depletes Arc from the postsynaptic terminal and prevents its activity-dependent recruitment. The combined use of gene-tagged and mutant mice allowed us to dissect the interactions of Arc with specific subsets of postsynaptic complexes. PSD95 supercomplexes are a family from which ~3% contain NMDA receptors (Frank et al., 2017b). The NMDA receptor requires PSD93 for

coassembly with PSD95 (Frank et al., 2016) and in the present study we found that PSD93 knockouts did not interfere with Arc-PSD95 interactions. Therefore, Arc can assemble with PSD95 supercomplexes that do not contain NMDA receptors. We also found that Arc did not interact with SAP102, which forms distinct complexes at ~350 kDa (Frank et al., 2016), nor did Arc require SAP102 for postsynaptic targeting. Together these results demonstrate that Arc is targeted to the postsynaptic terminal where it selectively interacts with signaling complexes organized by PSD95. Super-resolution microscopy has revealed that PSD95 and SAP102 are in separate nanodomains (Zheng et al., 2011) within the dendritic spine and that PSD95 nanodomains (Broadhead et al., 2016; Fukata et al., 2013; Nair et al., 2013) are positioned beneath the presynaptic release machinery (Tang et al., 2016). This suggests that Arc is selectively targeted by PSD95 to this critical region of the postsynaptic terminal where its supercomplexes participate in controlling synaptic transmission and plasticity.

Disruption of many proteins in Arc, PSD and many other proteins in the supercomplexes, lead to changes in synaptic plasticity and behavior, including knockout of Arc and PSD95 which both lead to enhanced LTP and impaired learning (Carlisle et al., 2008; Komiyama, submitted; Kopanitsa, submitted; Migaud et al., 1998; Nithianantharajah et al., 2013; Plath et al., 2006). A recent large-scale genetic screen of postsynaptic proteins in mice showed PSD95 supercomplexes were essential for the postsynaptic responses to simple and complex patterns of activity and the modulation of synaptic strength over a range of milliseconds to an hour (Kopanitsa et al., submitted). The supercomplexes were also essential for tuning the magnitude of innate and learned behavioral responses, including

simple and complex forms of behavior (Komiyama et al., submitted). Furthermore, these studies show that each innate and learned behavioral response required a specific subset or combination of postsynaptic proteins, which suggests that transient up-regulation and targeting of Arc to PSD95 supercomplexes will transiently modify behavior and synaptic physiology. This mechanism is consistent with the known role of Arc in learning.

The proteomes of the PSD and PSD95 supercomplexes are highly conserved between mice and humans (Bayes et al., 2011a) and specific genes (e.g. PSD93) have conserved roles in cognition (visuo-spatial learning, cognitive flexibility and attention) (Nithianantharajah et al., 2013). Our finding that human genetic disorders of cognition converge on Arc-PSD95 supercomplexes is in agreement with the mouse genetic findings. Here, wWe have reaffirmed the role of the supercomplexes in schizophrenia and extended the study to autism and intellectual disability. Moreover, the finding that variation in normal human intelligence and disorders of cognition involves the same sets of proteins indicates that genetic variation in Arc-PSD95 supercomplexes underpins the phenotypic continuum between normal cognitive variation and pathology.

There are over 130 brain diseases linked to mutations in the postsynaptic proteome (Bayes et al., 2011a) and a large number of uncharacterized multiprotein complexes (Frank et al., 2016), many of which contain at least one protein encoded by a disease gene. The integrated workflow shown here which is centered on genetically tagged mice and proteomic approaches offers a general and scalable approach toward understanding how the polygenic basis of brain disease is linked to the supramolecular organization of

proteins in the postsynaptic terminal of central synapses. All datasets are freely available through the Genes to Cognition website (www.genes2cognition.org).

EXPERIMENTAL PROCEDURES

Materials

Antibodies and reagents are detailed in Supplemental Experimental Procedures.

Animals

Arc transgenic mice (Arc^{TAP} and Arc-Venus), PSD95 knockout mice ($PSD95^{-/-}$) and the strains Arc^{TAP} and Arc-Venus crossed with $PSD95^{-/-}$ ($Arc^{TAP}xPSD95^{-/-}$, $Arc^{Venus}xPSD95^{-/-}$) used for experimental analysis were 2- to 5-month-old males from several intercrosses between the chimeras and C57BL/6 strain mice. Males were only used for experimental purposes. All animal experiments were conducted in a licensed animal facility in accordance to guidelines determined by the UK Animals (Scientific Procedures) Act, 1986 and all procedures were approved through the British Home Office Inspectorate. Animals were born and housed at the Wellcome Trust Sanger Institute, and exposed to conventional 12:12h light/dark cycles and *ad libitum* food supply. Mice were sacrificed by cervical dislocation and the forebrain was dissected on ice and snap-frozen in liquid nitrogen. Brain samples were stored at -80 ºC for few weeks prior to use.

See Supplemental Experimental Procedures for TAP and Venus Arc vector construction, gene targeting strategy and transgenic mice generation.

Tandem Affinity Purification

Tandem Affinity Purification (TAP) was perfomed as in (Fernandez et al., 2009). Briefly, mouse forebrain was homogenized on ice in 1% DOC buffer (50 mM Tris pH9.0, 1% Sodium Deoxycholate, 50 mM NaF, 20μM $ZnCl_2$, 1mM Na_3VO_4), 2 mM Pefabloc SC

(Roche) and 1 tablet/10ml protease inhibitor cocktail tablets (Roche) at 0.38 g wet weight per 7 ml cold buffer with a glass Teflon Douncer homogenizer. The homogenate was incubated for 1h at 4ºC and clarified at 50,000 *g* for 30 min at 4ºC. Isolation of Arc TAP tagged complexes was performed as described (Fernandez et al., 2009). The gel was fixed, stained with Colloidal Coomassie and lanes were cut into slices, destained and digested overnight with trypsin (Roche, Trypsin modified, sequencing grade) as described previously (Fernandez et al., 2009). Peptide digestion, LC-MS/MS, and proteomics data analysis is extensively described in Supplemental Experimental Procedures.

Blue Native PAGE

Blue-native PAGE has been performed as described in (Frank et al., 2016). See Supplemental Experimental Procedures for further description.

Cellular fractionation

Brain cellular fractionation was done according to the protocol described earlier (Chowdhury et al., 2006). See Supplemental Experimental Procedures for details.

Electrophysiology in hippocampal slices

Field excitatory postsynaptic potentials (fEPSPs) in the CA1 area of acute hippocampal slices were recorded using multi-electrode arrays as previously described (Coba et al., 2012; Kopanitsa et al., 2006). See Supplemental Experimental Procedures for detailed protocol.

Enrichment analysis of CNVs and rare coding mutations in Arc interactors in human neuropsychiatric disease.

In order to map rodent proteins onto human genes, all of the majority protein ids from Table S1 were first converted into both MGI & mouse NCBI/Entrez gene ids using the online ID mapping tool provided by Uniprot. These gene ids were then converted to human Entrez ids using the mapping file 'HOM_MouseHumanSequence.rpt', available from MGI (http://www.informatics.jax.org/). Any genes with a non-unique (e.g. 1-many) mapping between species, or where MGI and mouse Entrez ids mapped to different human genes, were excluded. De novo CNV enrichment analysis and de novo mutation exome sequencing datasets are detailed in Supplemental Experimental Procedures.

For the analysis of *de novo* mutations, we used DNENRICH, a freely-available software previously described (Fromer et al., 2014) (https://psychgen.u.hpc.mssm.edu/dnenrich/). Briefly, DNENRICH estimates the observed number of *de novo* mutations per gene or geneset by a genomic permutation strategy, which controls for gene size and structure, sequence coverage and local trinucleotide mutation rate. Significance of the enrichment in the rate and recurrence of mutations was assessed empirically by permutation. For the case/control data, we used the SMP algorithm (Purcell et al., 2014) which is part of the PLINK/Seq package (http://atgu.mgh.harvard.edu/plinkseq/). Briefly, the test for enrichment of a set of genes is based on the sum of gene-level case/control burden statistics relative to the exome-wide excess of burden in cases. Significance is assessed by permutation, comparing the observed distribution against 10,000 null replicates

(created by shuffling case/control labels within groups matched for ancestry and technical variables).

Neuropsychiatric genetic analyses: multiple testing correction

We first sought to verify the initial finding of Arc complex enrichment in *de novo* CNVs that has formed the foundation for all subsequent studies; this we correct for the 4 Arc complex gene-sets tested. We then turned to the analysis of rare coding variants. While there is consistent evidence for disease-variant enrichment across multiple independent studies (based on preliminary Arc data), these associations are modest. We therefore combine enrichment p-values from each rare-variant dataset to provide an overall summary of the evidence. These summary p-values are corrected for 12 tests: 4 Arc complex gene-sets tested x 2 classes of mutation (LoF and NS), plus the 4 *de novo* CNV tests already performed. This correction is likely to be conservative, as LoF and NS mutations are not independent sets. Finally, we present enrichment p-values for individual datasets, correcting for 52 tests (4 Arc complex gene-sets x LoF and NS mutations x 5 individual datasets, plus the 12 tests already performed).

Human cognitive ability phenotype and analysis

The phenotypes used in both the Cognitive Ageing in England and Scotland (CAGES) and the Brisbane Adolescent Twin Study (BATS) samples were taken from the summary data of (Hill et al., 2014a) and are described in Supplemental Experimental Procedures. Genome wide association had been carried out in each cohort of CAGES using Mach2QTL (Li et al., 2010) before being meta-analysed in METAL (Willer et al., 2010)

using an inverse variance weighted model. SNPs were then assigned to genes based on their position in the UCSC human genome browser hg 18 assembly with a 50kb boundary around each gene to capture any regulatory elements. A gene based statistic was derived using VEGAS (Liu et al., 2010b) to control for the number of SNPs assigned to each gene as well as patterns of linkage disequilibrium. In order to test the principal hypothesis that the genes of the Arc complex (See Table S13) will show a greater association, as a set, than those drawn from across the genome, each gene based p value was –log10 transformed and rank ordered. Gene Set Enrichment Analysis (GSEA) (Subramanian et al., 2005; Wang et al., 2007), a competitive test of enrichment, was then used to determine if the gene identifiers for the Arc gene set fall higher in the genome-wide rankings than would be expected by chance alone (Hill et al., 2014a; Hill et al., 2014b). This was done for each set by deriving a Kolmogorov-Smirnov (K-S) statistic weighted by the p-values of the gene based statistic in order to take into account both the ranks and the distance between ranks. Following this the genome wide ranked set was permuted and the K-S statistic calculated again. Statistical significance was established by using 15,000 permutations of the genome wide ranked set with the p-value describing the proportion of permuted K-S tests smaller than the original un-permuted K-S statistic. Statistical significance was set at < 0.05 and FDR <0.25 (Subramanian et al., 2005; Wang et al., 2007).

Immunostaining of brain sections and primary neurons, **forebrain extracts and western blot**, **PSD95 immunoprecipitation**, **classification of Arc interactors**

according to their gene ontology process, **creation of Arc interactor genesets**, **network building** and **statistics** are described in Supplemental Experimental Procedures.

SUPPLEMENTAL INFORMATION

Supplemental information includes Supplemental Experimental Procedures, 11 figures, 14 tables, Supplemental Figures and Table legends and Supplemental References.

ACCESSION NUMBERS

The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD007283.

AUTHOR CONTRIBUTIONS

Generation of knockin mice, EF, FZ, NHK; biochemistry, EF, RAWF; Immunohistochemistry, EF; mass spectrometry, MOC, JSC; ArcVenus imaging, JN, SAL, CM; electrophysiology, MVK; bioinformatics, MDRC, CM, JDA; disease genetics, AJP, MF, SMP; cognition genetics, WDH, IJD; technical assistance, DF, KAE, CLM, GC; direction, CB, SGNG; writing, EF, AJP, SGNG; conception and management, SGNG.

ACKNOWLEDGEMENTS

~~We thank K Elsegood for mouse colony management, C McLaughlin for tissue culture, D Maizels for artwork and Giulia Cencelli for help on Western blotting.~~ The research leading to these results has received funding from the European Union Seventh Framework

Formatted: Level 1

Formatted: Font: Bold

Programme under grant agreement number HEALTH-F2-2009-242167 (“SynSys-project”) and HEALTH-F2-2009-241498 (“EUROSPIN” project) and the Wellcome Trust. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. The work on the human cognitive ability was undertaken by The University of Edinburgh Centre for Cognitive Ageing and Cognitive Epidemiology (CCACE), part of the cross council Lifelong Health and Wellbeing Initiative (MR/K026992/1). EF was supported by a Federation of European Biochemistry Societies postdoctoral fellowship, Marie Curie Intra European fellowship and Department of Defense, USA. Funding from the Biotechnology and Biological Sciences Research Council (BBSRC) and Medical Research Council (MRC) is gratefully acknowledged. Ian J. Deary (IJD) is supported by CCACE. Claudia Bagni is supported by FWO, Department of Defense USA and Associazione Italiana Sindrome X Fragile. W David Hill is supported by a grant from Age UK (the Disconnected Ming project) to IJD. D Maizels for artwork.

The authors declare no competing interests.

FIGURES AND TABLE LEGENDS

Figure 1. Tandem Affinity Purification of Arc reveals postsynaptic complexes with a native size of 1.5 MDa. (A) Structure of the TAP tagged Arc regions including a potential coiled coil domain, a SH3-endophilin-2 and -3 binding region, a dynamin-2 binding region and the C-terminal TAP tag sequence domain before the stop codon of the protein. **(B)** BNP of wild type and $Arc^{TAP/+}$ forebrain extracts blotted with Flag and Arc antibodies where Arc and TAP tagged Arc can be mainly detected at 1.5 megadalton (MDa). Closed arrow indicates the main Arc complex whereas open arrows indicate lower molecular weight Arc complexes. **(C)** Arc was tandem affinity purified from $Arc^{TAP/TAP}$ forebrain extracts, eluted and collected in 5 consecutive fractions following HAT purification. **(D)** Colloidal Coomassie staining of three independent tandem affinity purifications from wild type (left) and $Arc^{TAP/TAP}$ (right) forebrains. The lanes were cut for LC-MS/MS and the identified proteins are listed in Table S1. Arc and the Tev enzyme are indicated. **(E)** The absolute expression value of each protein in the tandem purification was estimated by the iBAQ intensity values obtained in each purification. **(F)** BNP of wild type, $Arc^{TAP/+}$ and $PSD95^{TAP/+}$ forebrain extracts blotted with Flag antibody. TAP tagged Arc and TAP tagged PSD95 levels are detected. WT, wild type; TAP/+, heterozygous for TAP tagged Arc or PSD95 as indicated; SN: supernatant; Mw: molecular weight; MDa: megadalton; kDa: kilodalton.

Figure 2. Arc protein levels are reduced in fractions of PSD95 knockout mice. (**A**) Representative immunoblot showing relative abundance of PSD95, Arc, GluR1 and GluR2 proteins in total hippocampal lysates from PSD95 knockout ($PSD95^{-/-}$) and

matched WT littermates. Arc is reduced to 35.0 ± 17.1% of the wild type in the PSD95 mutant mice (N=4 for each matched pair *P<0.05). nNOS was used as loading control. **(B)** Representative Arc staining of sagittal sections of hippocampus (left panels) and magnification of the granular layer (right panels) for wild type and PSD95 knockout mice. *DG*: Dentate gyrus. Bar = 1mm. **(C)** Hippocampal extracts of PSD95 mutant and WT mutant mice were biochemically fractionated into synaptosomes and into cytoskeletal and vesicular components, referred to as "Light". The synaptosomal fraction was subsequently dissociated into PSDs and Triton X-100 soluble fraction. Arc levels were dramatically reduced in the PSD95 mutant while no changes in GluR1, GluN2B, IRSp53 and Rac1 proteins were observed. **(D)** BNP of WT and PSD95 knockout forebrain extracts blotted with Arc and PSD95 antibodies. Long exposure of the blots shows Arc complexes migrating at a lower molecular weight than 1.20 MDa (middle panel). **(E)** Representative section of Arc^{Venus} brain crossed with WT (left panel) and $PSD95^{-/-}$ (right panel) mice. Bar chart of the total cell fluorescence corrected by the area and the background signal. Bar = 15 μm. WT: wild type; $PSD95^{-/-}$: PSD95 knockout mice; SN: Triton X-100 soluble fraction. Mw: molecular weight; MDa: megadalton; kDa: kilodalton. BNP: Blue Native PAGE.

Figure 3. Quantitative proteomics analysis of Arc^{TAP} reveals a depletion of postsynaptic proteins in the PSD95 knockout mice. (A) Dimethyl labeling-based quantitative mass spectrometry of TAP purified proteins from $Arc^{TAP/+}$ and $Arc^{TAP/TAP}$ crossed with PSD95 knockout forebrain ($Arc^{TAP/TAP}PSD95^{-/-}$). Plot displays enrichment ratios of $Arc^{TAP/TAP}PSD95^{-/-}$ *versus* $Arc^{TAP/+}$ (x axis) and iBAQ enrichment values of the

step purification (y axis). Proteins meeting criteria for enrichment (>1.5 fold) are highlighted in green and for depletion (<0.667 fold) are highlighted in red. Name of depleted and enriched PSD95 interactors is indicated. See Supplemental Experimental Procedures for enrichment criteria. **(B)** Mouse interactome network constructed from the publicly available databases BioGrid, DIP, IntAct, MINT, STRING, UniProt, BIND and mentha using the Psicquic software package. Network is visualised using Visone. Proteins highlighted green/red meet the enrichment/depletion criteria discussed in the Experimental procedures.

Table I. Arc gene set analysis of autism, schizophrenia, epilepsy, intellectual disability (ID) and schizophrenia candidate gene sets. Enrichment test empirical *P* values for autism (De Rubeis et al., 2014; Iossifov et al., 2012; Jiang et al., 2013), epilepsy (Euro et al., 2014), ID (de Ligt et al., 2012; Hamdan et al., 2014; Rauch et al., 2012) and schizophrenia (Fromer et al., 2014; Girard et al., 2011; Gulsuner et al., 2013; Kirov et al., 2012; McCarthy et al., 2014; Xu et al., 2012). PSD95$^{-/-}$: PSD95/*Dlg4* knockout mice**.**

REFERENCES

Babu, M., Arnold, R., Bundalovic-Torma, C., Gagarinova, A., Wong, K.S., Kumar, A., Stewart, G., Samanfar, B., Aoki, H., Wagih, O.*, et al.* (2014). Quantitative genome-wide genetic interaction screens reveal global epistatic relationships of protein complexes in Escherichia coli. PLoS genetics *10*, e1004120.

Bah, A., Vernon, R.M., Siddiqui, Z., Krzeminski, M., Muhandiram, R., Zhao, C., Sonenberg, N., Kay, L.E., and Forman-Kay, J.D. (2015). Folding of an intrinsically disordered protein by phosphorylation as a regulatory switch. Nature *519*, 106-109.

Bayes, A., Collins, M.O., Croning, M.D., van de Lagemaat, L.N., Choudhary, J.S., and Grant, S.G. (2012). Comparative study of human and mouse postsynaptic proteomes finds high compositional conservation and abundance differences for key synaptic proteins. PLoS One *7*, e46683.

Bayes, A., Collins, M.O., Galtrey, C.M., Simonnet, C., Roy, M., Croning, M.D., Gou, G., van de Lagemaat, L.N., Milward, D., Whittle, I.R.*, et al.* (2014). Human post-mortem synapse proteome integrity screening for proteomic studies of postsynaptic complexes. Molecular brain *7*, 88.

Bayes, A., Collins, M.O., Reig-Viader, R., Gou, G., Goulding, D., Izquierdo, A., Choudhary, J.S., Emes, R.D., and Grant, S.G. (2017). Evolution of complexity in the zebrafish synapse proteome. Nature communications *8*, 14613.

Bayes, A., van de Lagemaat, L.N., Collins, M.O., Croning, M.D., Whittle, I.R., Choudhary, J.S., and Grant, S.G. (2011a). Characterization of the proteome, diseases and evolution of the human postsynaptic density. Nature neuroscience *14*, 19-21.

Bayes, A., van de Lagemaat, L.N., Collins, M.O., Croning, M.D., Whittle, I.R., Choudhary, J.S., and Grant, S.G. (2011b). Characterization of the proteome, diseases and evolution of the human postsynaptic density. Nature neuroscience *14*, 19-21.

Brandes, U., and Wagner, D., (2012). Visone, Analysis and visualization of social networks, version 2.6.5.

Broadhead, M.J., Horrocks, M.H., Zhu, F., Muresan, L., Benavides-Piccione, R., DeFelipe, J., Fricker, D., Kopanitsa, M.V., Duncan, R.R., Klenerman, D.*, et al.* (2016). PSD95 nanoclusters are postsynaptic building blocks in hippocampus circuits. Sci Rep *6*, 24626.

Brose, N., O'Connor, V., and Skehel, P. (2010). Synaptopathy: dysfunction of synaptic function? Biochemical Society transactions *38*, 443-444.

Carlisle, H.J., Fink, A.E., Grant, S.G., and O'Dell, T.J. (2008). Opposing effects of PSD-93 and PSD-95 on long-term potentiation and spike timing-dependent plasticity. The Journal of physiology *586*, 5885-5900.

Chowdhury, S., Shepherd, J.D., Okuno, H., Lyford, G., Petralia, R.S., Plath, N., Kuhl, D., Huganir, R.L., and Worley, P.F. (2006). Arc/Arg3.1 interacts with the endocytic machinery to regulate AMPA receptor trafficking. Neuron *52*, 445-459.

Coba, M.P., Komiyama, N.H., Nithianantharajah, J., Kopanitsa, M.V., Indersmitten, T., Skene, N.G., Tuck, E.J., Fricker, D.G., Elsegood, K.A., Stanford, L.E.*, et al.* (2012). TNiK is required for postsynaptic and nuclear signaling pathways and cognitive function. The Journal of neuroscience : the official journal of the Society for Neuroscience *32*, 13987-13999.

Daberkow, D.P., Riedy, M.D., Kesner, R.P., and Keefe, K.A. (2007). Arc mRNA induction in striatal efferent neurons associated with response learning. The European journal of neuroscience *26*, 228-241.

Davies, G., Tenesa, A., Payton, A., Yang, J., Harris, S.E., Liewald, D., Ke, X., Le Hellard, S., Christoforou, A., Luciano, M.*, et al.* (2011). Genome-wide association studies establish that human intelligence is highly heritable and polygenic. Mol Psychiatry *16*, 996-1005.

de Ligt, J., Willemsen, M.H., van Bon, B.W., Kleefstra, T., Yntema, H.G., Kroes, T., Vulto-van Silfhout, A.T., Koolen, D.A., de Vries, P., Gilissen, C.*, et al.* (2012). Diagnostic exome sequencing in persons with severe intellectual disability. The New England journal of medicine *367*, 1921-1929.

De Rubeis, S., He, X., Goldberg, A.P., Poultney, C.S., Samocha, K., Cicek, A.E., Kou, Y., Liu, L., Fromer, M., Walker, S.*, et al.* (2014). Synaptic, transcriptional and chromatin genes disrupted in autism. Nature *515*, 209-215.

de Zubicaray, G.I., Chiang, M.C., McMahon, K.L., Shattuck, D.W., Toga, A.W., Martin, N.G., Wright, M.J., and Thompson, P.M. (2008). Meeting the Challenges of Neuroimaging Genetics. Brain Imaging Behav *2*, 258-263.

Deary, I.J., Gow, A.J., Pattie, A., and Starr, J.M. (2012). Cohort profile: the lothian birth cohorts of 1921 and 1936. Int J Epidemiol *41*, 1576-1584.

Distler, U., Schmeisser, M.J., Pelosi, A., Reim, D., Kuharev, J., Weiczner, R., Baumgart, J., Boeckers, T.M., Nitsch, R., Vogt, J.*, et al.* (2014). In-depth protein profiling of the postsynaptic density from mouse hippocampus using data-independent acquisition proteomics. Proteomics *14*, 2607-2613.

Euro, E.-R.E.S.C., Epilepsy Phenome/Genome, P., and Epi, K.C. (2014). De novo mutations in synaptic transmission genes including DNM1 cause epileptic encephalopathies. American journal of human genetics *95*, 360-370.

Fernandez, E., Collins, M.O., Uren, R.T., Kopanitsa, M.V., Komiyama, N.H., Croning, M.D., Zografos, L., Armstrong, J.D., Choudhary, J.S., and Grant, S.G. (2009). Targeted tandem affinity purification of PSD-95 recovers core postsynaptic complexes and schizophrenia susceptibility proteins. Molecular systems biology *5*, 269.

Fitzgerald, P.J., Pinard, C.R., Camp, M.C., Feyder, M., Sah, A., Bergstrom, H.C., Graybeal, C., Liu, Y., Schluter, O.M., Grant, S.G.*, et al.* (2014). Durable fear memories require PSD-95. Molecular psychiatry.

Frank, R.A., and Grant, S.G. (2017). Supramolecular organization of NMDA receptors and the postsynaptic density. Current opinion in neurobiology *45*, 139-147.

Frank, R.A., Komiyama, N.H., Ryan, T.J., Zhu, F., O'Dell, T.J., and Grant, S.G. (2016). NMDA receptors are selectively partitioned into complexes and supercomplexes during synapse maturation. Nature communications *7*, 11264.

Frank, R.A., Zhu, F., Komiyama, N.H., and Grant, S.G. (2017a). Hierarchical organisation and genetically separable subfamilies of PSD95 postsynaptic supercomplexes. Journal of Neurochemistry (submitted).

Frank, R.A.W., Zhu, F., Komiyama, N.H., and Grant, S.G.N. (2017b). Hierarchical organisation and genetically separable subfamilies of PSD95 postsynaptic supercomplexes. J Neurochem.

Fromer, M., Pocklington, A.J., Kavanagh, D.H., Williams, H.J., Dwyer, S., Gormley, P., Georgieva, L., Rees, E., Palta, P., Ruderfer, D.M.*, et al.* (2014). De novo mutations in schizophrenia implicate synaptic networks. Nature *506*, 179-184.

Fukata, Y., Dimitrov, A., Boncompain, G., Vielemeyer, O., Perez, F., and Fukata, M. (2013). Local palmitoylation cycles define activity-regulated postsynaptic subdomains. The Journal of cell biology *202*, 145-161.

Gao, M., Sossa, K., Song, L., Errington, L., Cummings, L., Hwang, H., Kuhl, D., Worley, P., and Lee, H.K. (2010). A specific requirement of Arc/Arg3.1 for visual experience-induced homeostatic synaptic plasticity in mouse primary visual cortex. The Journal of neuroscience : the official journal of the Society for Neuroscience *30*, 7168-7178.

Gibbs, E.B., Lu, F., Portz, B., Fisher, M.J., Medellin, B.P., Laremore, T.N., Zhang, Y.J., Gilmour, D.S., and Showalter, S.A. (2017). Phosphorylation induces sequence-specific conformational switches in the RNA polymerase II C-terminal domain. Nature communications *8*, 15233.

Girard, S.L., Gauthier, J., Noreau, A., Xiong, L., Zhou, S., Jouan, L., Dionne-Laporte, A., Spiegelman, D., Henrion, E., Diallo, O.*, et al.* (2011). Increased exonic de novo mutation rate in individuals with schizophrenia. Nature genetics *43*, 860-863.

Grant, S.G. (2012). Synaptopathies: diseases of the synaptome. Curr Opin Neurobiol *22*, 522-529.

Grant, S.G., Marshall, M.C., Page, K.L., Cumiskey, M.A., and Armstrong, J.D. (2005). Synapse proteomics of multiprotein complexes: en route from genes to nervous system diseases. Hum Mol Genet *14 Spec No. 2*, R225-234.

Grant, S.G.N. (2013). Synaptic Disease in Psychiatry. In Neurobiology of Mental Illness, J.D.B. Dennis S. Charney, Pamela Sklar, E. J. Nestler, ed. (Oxford: Oxford University Press).

Greer, P.L., Hanayama, R., Bloodgood, B.L., Mardinly, A.R., Lipton, D.M., Flavell, S.W., Kim, T.K., Griffith, E.C., Waldon, Z., Maehr, R.*, et al.* (2010). The Angelman Syndrome protein Ube3A regulates synapse development by ubiquitinating arc. Cell *140*, 704-716.

Gulsuner, S., Walsh, T., Watts, A.C., Lee, M.K., Thornton, A.M., Casadei, S., Rippey, C., Shahin, H., Nimgaonkar, V.L., Go, R.C.*, et al.* (2013). Spatial and temporal mapping of de novo mutations in schizophrenia to a fetal prefrontal cortical network. Cell *154*, 518-529.

Guzowski, J.F., Lyford, G.L., Stevenson, G.D., Houston, F.P., McGaugh, J.L., Worley, P.F., and Barnes, C.A. (2000). Inhibition of activity-dependent arc protein expression in the rat hippocampus impairs the maintenance of long-term potentiation and the consolidation of long-term memory. The Journal of neuroscience : the official journal of the Society for Neuroscience *20*, 3993-4001.

Guzowski, J.F., McNaughton, B.L., Barnes, C.A., and Worley, P.F. (1999). Environment-specific expression of the immediate-early gene Arc in hippocampal neuronal ensembles. Nature neuroscience *2*, 1120-1124.

Hamdan, F.F., Srour, M., Capo-Chichi, J.M., Daoud, H., Nassif, C., Patry, L., Massicotte, C., Ambalavanan, A., Spiegelman, D., Diallo, O.*, et al.* (2014). De novo mutations in moderate or severe intellectual disability. PLoS genetics *10*, e1004772.

Hill, W.D., Davies, G., van de Lagemaat, L.N., Christoforou, A., Marioni, R.E., Fernandes, C.P.D., Liewald, D.C., Croning, M.D.R., Payton, A., Craig, L.C.A.*, et al.* (2014a). Human cognitive ability is influenced by genetic variation in components of postsynaptic signalling complexes assembled by NMDA receptors and MAGUK proteins. Transl Psychiatry *4*, e341.

Hill, W.D., de Leeuw, C., Davies, G., Liewald, D.C.M., Payton, A., Craig, L.C., J.W., L., Horan, M., Ollier, W., Starr, J.M.*, et al.* (2014b). Functional Gene Group Analysis Indicates No Role for Heterotrimeric G Proteins in Cognitive Ability. PloS one *9*.

Husi, H., and Grant, S.G. (2001). Isolation of 2000-kDa complexes of N-methyl-D-aspartate receptor and postsynaptic density 95 from mouse brain. J Neurochem *77*, 281-291.

Husi, H., Ward, M.A., Choudhary, J.S., Blackstock, W.P., and Grant, S.G. (2000). Proteomic analysis of NMDA receptor-adhesion protein signaling complexes. Nature neuroscience *3*, 661-669.

Iossifov, I., Ronemus, M., Levy, D., Wang, Z., Hakker, I., Rosenbaum, J., Yamrom, B., Lee, Y.H., Narzisi, G., Leotta, A.*, et al.* (2012). De novo gene disruptions in children on the autistic spectrum. Neuron *74*, 285-299.

Jakkamsetti, V., Tsai, N.P., Gross, C., Molinaro, G., Collins, K.A., Nicoletti, F., Wang, K.H., Osten, P., Bassell, G.J., Gibson, J.R.*, et al.* (2013). Experience-induced Arc/Arg3.1 primes CA1 pyramidal neurons for metabotropic glutamate receptor-dependent long-term synaptic depression. Neuron *80*, 72-79.

Jiang, Y.H., Yuen, R.K., Jin, X., Wang, M., Chen, N., Wu, X., Ju, J., Mei, J., Shi, Y., He, M.*, et al.* (2013). Detection of clinically relevant genetic variants in autism spectrum disorder by whole-genome sequencing. American journal of human genetics *93*, 249-263.

Kelly, M.P., and Deadwyler, S.A. (2003). Experience-dependent regulation of the immediate-early gene arc differs across brain regions. The Journal of neuroscience : the official journal of the Society for Neuroscience *23*, 6443-6451.

Kerrigan, T.L., and Randall, A.D. (2013). A new player in the "synaptopathy" of Alzheimer's disease - arc/arg 3.1. Frontiers in neurology *4*, 9.

Kirov, G., Pocklington, A.J., Holmans, P., Ivanov, D., Ikeda, M., Ruderfer, D., Moran, J., Chambert, K., Toncheva, D., Georgieva, L.*, et al.* (2012). De novo CNV analysis implicates specific abnormalities of postsynaptic signalling complexes in the pathogenesis of schizophrenia. Molecular psychiatry *17*, 142-153.

Komiyama, N.H., van de Lagemaat, L.N., Stanford, L.E., Pettit, C., Strathdee, D.J., Strathdee, J.E., Fricker, D.G., Tuck, E.J., Elsegood, K.A., Ryan, T.J., et al. (submitted). Large-scale genetic dissection of postsynaptic mechanisms controlling vertebrate behavioral repertoires.

Komiyama, N.H., Watabe, A.M., Carlisle, H.J., Porter, K., Charlesworth, P., Monti, J., Strathdee, D.J., O'Carroll, C.M., Martin, S.J., Morris, R.G.*, et al.* (2002). SynGAP regulates ERK/MAPK signaling, synaptic plasticity, and learning in the complex with postsynaptic density 95 and NMDA receptor. J Neurosci *22*, 9721-9732.

Kopanitsa, M.V., Afinowi, N.O., and Grant, S.G. (2006). Recording long-term potentiation of synaptic transmission by three-dimensional multi-electrode arrays. BMC neuroscience *7*, 61.

Kopanitsa, M.V., van de Lagemaat, L.N., Afinowi, N.O., Strathdee, D.J., Strathdee, K.E., Fricker, D.G., Tuck, E.J., Elsegood, K.A., Croning, M.D., Komiyama, N.H.*, et al.* (submitted). A mouse genetic screen of postsynaptic proteins controlling synaptic responses to patterns of neural activity.

Korb, E., Wilkinson, C.L., Delgado, R.N., Lovero, K.L., and Finkbeiner, S. (2013). Arc in the nucleus regulates PML-dependent GluA1 transcription and homeostatic plasticity. Nature neuroscience *16*, 874-883.

Li, L., Carter, J., Gao, X., Whitehead, J., and Tourtellotte, W.G. (2005). The neuroplasticity-associated arc gene is a direct transcriptional target of early growth response (Egr) transcription factors. Molecular and cellular biology *25*, 10286-10300.

Li, Y., Willer, C.J., Ding, J., Scheet, P., and Abecasis, G.R. (2010). MaCH: using sequence and genotype data to estimate haplotypes and unobserved genotypes. Genet Epidemiol *34*, 816-834.

Link, W., Konietzko, U., Kauselmann, G., Krug, M., Schwanke, B., Frey, U., and Kuhl, D. (1995). Somatodendritic expression of an immediate early gene is regulated by synaptic activity. Proceedings of the National Academy of Sciences of the United States of America *92*, 5734-5738.

Liu, J.Z., McRae, A.F., Nyholt, D.R., Medland, S.E., Wray, N.R., Brown, K.M., Hayward, N.K., Montgomery, G.W., Visscher, P.M., Martin, N.G.*, et al.* (2010a). A versatile gene-based test for genome-wide association studies. Am J Hum Genet *87*, 139-145.

Liu, J.Z., McRae, A.F., Nyholt, D.R., Medland, S.E., Wray, N.R., Brown, K.M., Hayward, N.K., Montgomery, G.W., Visscher, P.M., Martin, N.G.*, et al.* (2010b). A versatile gene-based test for genome-wide association studies. American journal of human genetics *87*, 139-145.

Lu, X., Kensche, P.R., Huynen, M.A., and Notebaart, R.A. (2013). Genome evolution predicts genetic interactions in protein complexes and reveals cancer drug targets. Nature communications *4*, 2124.

Lyford, G.L., Yamagata, K., Kaufmann, W.E., Barnes, C.A., Sanders, L.K., Copeland, N.G., Gilbert, D.J., Jenkins, N.A., Lanahan, A.A., and Worley, P.F. (1995). Arc, a growth factor and activity-regulated gene, encodes a novel cytoskeleton-associated protein that is enriched in neuronal dendrites. Neuron *14*, 433-445.

Manago, F., Mereu, M., Mastwal, S., Mastrogiacomo, R., Scheggia, D., Emanuele, M., De Luca, M.A., Weinberger, D.R., Wang, K.H., and Papaleo, F. (2016). Genetic

Disruption of Arc/Arg3.1 in Mice Causes Alterations in Dopamine and Neurobehavioral Phenotypes Related to Schizophrenia. Cell reports *16*, 2116-2128.

McCarthy, S.E., Gillis, J., Kramer, M., Lihm, J., Yoon, S., Berstein, Y., Mistry, M., Pavlidis, P., Solomon, R., Ghiban, E.*, et al.* (2014). De novo mutations in schizophrenia implicate chromatin remodeling and support a genetic overlap with autism and intellectual disability. Molecular psychiatry *19*, 652-658.

McCurry, C.L., Shepherd, J.D., Tropea, D., Wang, K.H., Bear, M.F., and Sur, M. (2010). Loss of Arc renders the visual cortex impervious to the effects of sensory experience or deprivation. Nature neuroscience *13*, 450-457.

Migaud, M., Charlesworth, P., Dempster, M., Webster, L.C., Watabe, A.M., Makhinson, M., He, Y., Ramsay, M.F., Morris, R.G., Morrison, J.H.*, et al.* (1998). Enhanced long-term potentiation and impaired learning in mice with mutant postsynaptic density-95 protein. Nature *396*, 433-439.

Miyashita, T., Kubik, S., Haghighi, N., Steward, O., and Guzowski, J.F. (2009). Rapid activation of plasticity-associated gene transcription in hippocampal neurons provides a mechanism for encoding of one-trial experience. The Journal of neuroscience : the official journal of the Society for Neuroscience *29*, 898-906.

Moga, D.E., Calhoun, M.E., Chowdhury, A., Worley, P., Morrison, J.H., and Shapiro, M.L. (2004). Activity-regulated cytoskeletal-associated protein is localized to recently activated excitatory synapses. Neuroscience *125*, 7-11.

Nair, D., Hosy, E., Petersen, J.D., Constals, A., Giannone, G., Choquet, D., and Sibarita, J.B. (2013). Super-resolution imaging reveals that AMPA receptors inside synapses are dynamically organized in nanodomains regulated by PSD95. The Journal of neuroscience : the official journal of the Society for Neuroscience *33*, 13204-13224.

Nithianantharajah, J., Komiyama, N.H., McKechanie, A., Johnstone, M., Blackwood, D.H., St Clair, D., Emes, R.D., van de Lagemaat, L.N., Saksida, L.M., Bussey, T.J.*, et al.* (2013). Synaptic scaffold evolution generated components of vertebrate cognitive complexity. Nature neuroscience *16*, 16-24.

Okuno, H., Akashi, K., Ishii, Y., Yagishita-Kyo, N., Suzuki, K., Nonaka, M., Kawashima, T., Fujii, H., Takemoto-Kimura, S., Abe, M.*, et al.* (2012). Inverse synaptic tagging of inactive synapses via dynamic interaction of Arc/Arg3.1 with CaMKIIbeta. Cell *149*, 886-898.

Park, S., Park, J.M., Kim, S., Kim, J.A., Shepherd, J.D., Smith-Hicks, C.L., Chowdhury, S., Kaufmann, W., Kuhl, D., Ryazanov, A.G.*, et al.* (2008). Elongation factor 2 and fragile X mental retardation protein control the dynamic translation of Arc/Arg3.1 essential for mGluR-LTD. Neuron *59*, 70-83.

Plath, N., Ohana, O., Dammermann, B., Errington, M.L., Schmitz, D., Gross, C., Mao, X., Engelsberg, A., Mahlke, C., Welzl, H.*, et al.* (2006). Arc/Arg3.1 is essential for the consolidation of synaptic plasticity and memories. Neuron *52*, 437-444.

Pocklington, A.J., Cumiskey, M., Armstrong, J.D., and Grant, S.G. (2006). The proteomes of neurotransmitter receptor complexes form modular networks with distributed functionality underlying plasticity and behaviour. Molecular systems biology *2*, 2006 0023.

Purcell, S.M., Moran, J.L., Fromer, M., Ruderfer, D., Solovieff, N., Roussos, P., O'Dushlaine, C., Chambert, K., Bergen, S.E., Kahler, A.*, et al.* (2014). A polygenic burden of rare disruptive mutations in schizophrenia. Nature *506*, 185-190.

Rabbitt, P.M.A., McInnes, L., Diggle, P., Holland, F., Bent, N., Abson, V., Pendleton, N., and Horan, M. (2004). The University of Manchester Longitudinal Study of Cognition in Normal Healthy Old Age, 1983 through 2003. Neuropsychol Dev Cogn B Aging Neuropsychol Cogn *11*, 245-279.

Rauch, A., Wieczorek, D., Graf, E., Wieland, T., Endele, S., Schwarzmayr, T., Albrecht, B., Bartholdi, D., Beygo, J., Di Donato, N.*, et al.* (2012). Range of genetic mutations associated with severe non-syndromic sporadic intellectual disability: an exome sequencing study. Lancet *380*, 1674-1682.

Rial Verde, E.M., Lee-Osbourne, J., Worley, P.F., Malinow, R., and Cline, H.T. (2006). Increased expression of the immediate-early gene arc/arg3.1 reduces AMPA receptor-mediated synaptic transmission. Neuron *52*, 461-474.

Ryan, T.J., Emes, R.D., Grant, S.G., and Komiyama, N.H. (2008). Evolution of NMDA receptor cytoplasmic interaction domains: implications for organisation of synaptic signalling complexes. BMC neuroscience *9*, 6.

Ryan, T.J., Kopanitsa, M.V., Indersmitten, T., Nithianantharajah, J., Afinowi, N.O., Pettit, C., Stanford, L.E., Sprengel, R., Saksida, L.M., Bussey, T.J.*, et al.* (2013). Evolution of GluN2A/B cytoplasmic domains diversified vertebrate synaptic plasticity and behavior. Nature neuroscience *16*, 25-32.

Schwanhausser, B., Busse, D., Li, N., Dittmar, G., Schuchhardt, J., Wolf, J., Chen, W., and Selbach, M. (2011). Global quantification of mammalian gene expression control. Nature *473*, 337-342.

Shepherd, J.D., Rumbaugh, G., Wu, J., Chowdhury, S., Plath, N., Kuhl, D., Huganir, R.L., and Worley, P.F. (2006). Arc/Arg3.1 mediates homeostatic synaptic scaling of AMPA receptors. Neuron *52*, 475-484.

Steward, O., Wallace, C.S., Lyford, G.L., and Worley, P.F. (1998). Synaptic activation causes the mRNA for the IEG Arc to localize selectively near activated postsynaptic sites on dendrites. Neuron *21*, 741-751.

Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., Paulovich, A., Pomeroy, S.L., Golub, T.R., Lander, E.S., *et al.* (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A *102*, 15545-15550.

Tang, A.H., Chen, H., Li, T.P., Metzbower, S.R., MacGillavry, H.D., and Blanpied, T.A. (2016). A trans-synaptic nanocolumn aligns neurotransmitter release to receptors. Nature *536*, 210-214.

Vazdarjanova, A., and Guzowski, J.F. (2004). Differences in hippocampal neuronal population responses to modifications of an environmental context: evidence for distinct, yet complementary, functions of CA3 and CA1 ensembles. The Journal of neuroscience : the official journal of the Society for Neuroscience *24*, 6489-6496.

Vazdarjanova, A., Ramirez-Amaya, V., Insel, N., Plummer, T.K., Rosi, S., Chowdhury, S., Mikhael, D., Worley, P.F., Guzowski, J.F., and Barnes, C.A. (2006). Spatial exploration induces ARC, a plasticity-related immediate-early gene, only in calcium/calmodulin-dependent protein kinase II-positive principal excitatory and inhibitory neurons of the rat forebrain. The Journal of comparative neurology *498*, 317-329.

Vidal, M., Cusick, M.E., and Barabasi, A.L. (2011). Interactome networks and human disease. Cell *144*, 986-998.

Wang, K., Li, M., and Bucan, M. (2007). Pathway-based approaches for analysis of genomewide association studies. American journal of human genetics *81*, 1278-1283.
Wang, K.H., Majewska, A., Schummers, J., Farley, B., Hu, C., Sur, M., and Tonegawa, S. (2006). In vivo two-photon imaging reveals a role of arc in enhancing orientation specificity in visual cortex. Cell *126*, 389-402.

Waung, M.W., Pfeiffer, B.E., Nosyreva, E.D., Ronesi, J.A., and Huber, K.M. (2008). Rapid translation of Arc/Arg3.1 selectively mediates mGluR-dependent LTD through persistent increases in AMPAR endocytosis rate. Neuron *59*, 84-97.

Whalley, L.J., Murray, A.D., Staff, R.T., Starr, J.M., Deary, I.J., Fox, H.C., Lemmon, H., Duthie, S.J., Collins, A.R., and Crawford, J.R. (2011). How the 1932 and 1947 mental surveys of Aberdeen schoolchildren provide a framework to explore the childhood origins of late onset disease and disability. Maturitas *69*, 365-372.

Wibrand, K., Pai, B., Siripornmongcolchai, T., Bittins, M., Berentsen, B., Ofte, M.L., Weigel, A., Skaftnesmo, K.O., and Bramham, C.R. (2012). MicroRNA regulation of the synaptic plasticity-related gene Arc. PloS one *7*, e41688.

Willer, C.J., Li, Y., and Abecasis, G.R. (2010). METAL: fast and efficient meta-analysis of genomewide association scans. Bioinformatics *26*, 2190-2191.

Wright, M.J., De Geus, E., Ando, J., Luciano, M., Posthuma, D., Ono, Y., Hansell, N., Van Baal C, Hiraishi, K., Hasegawa, T.*, et al.* (2001). Genetics of cognition: outline of a collaborative twin study. Twin Research *4*, 48-56.

Wright, M.J., and Martin, N., G. (2004). The Brisbane Adolescent Twin Study: outline of study methods and research projects. Australian Journal of Psychology *56*, 65-78.

Xu, B., Ionita-Laza, I., Roos, J.L., Boone, B., Woodrick, S., Sun, Y., Levy, S., Gogos, J.A., and Karayiorgou, M. (2012). De novo gene mutations highlight patterns of genetic and neural complexity in schizophrenia. Nature genetics *44*, 1365-1369.

Xue, B., Dunbrack, R.L., Williams, R.W., Dunker, A.K., and Uversky, V.N. (2010). PONDR-FIT: a meta-predictor of intrinsically disordered amino acids. Biochim Biophys Acta *1804*, 996-1010.

Zhang, W., Wu, J., Ward, M.D., Yang, S., Chuang, Y.A., Xiao, M., Li, R., Leahy, D.J., and Worley, P.F. (2015). Structural basis of arc binding to synaptic proteins: implications for cognitive disease. Neuron *86*, 490-500.

Zheng, C.Y., Wang, Y.X., Kachar, B., and Petralia, R.S. (2011). Differential localization of SAP102 and PSD-95 is revealed in hippocampal spines using super-resolution light microscopy. Commun Integr Biol *4*, 104-105.