Invited Review

Gene expression in maturing neurons: regulatory mechanisms and related neurodevelopmental disorders

DING Baojin*
Department of Neurobiology, University of Massachusetts Medical School, Worcester, MA 01655, USA

Abstract: During the central nervous system (CNS) development, the interactions between intrinsic genes and extrinsic environment ensure that each neuronal developmental stage (e.g., neuronal proliferation, differentiation, migration, axon extension, dendritogenesis and formation of functional synapses) occurs in the proper timing and sequence. The successful coordination requires that numerous groups of genes are exquisitely regulated in a spatiotemporal manner by various regulatory mechanisms, including sequence-specific DNA-binding proteins, histone modifications, DNA methylation, chromatin remodeling, and microRNAs (miRNAs). By targeting chromatin structure, transcription and translation processes, these mechanisms form a regulatory network to accomplish the fine regulation of gene expression in response to environmental stimuli at different developmental stages. Dysregulation of the gene expression during neuronal development has been shown to be implicated in a number of neurodevelopmental disorders, such as autism spectrum disorders (ASD), Rett syndrome (RTT), Fragile-X syndrome (FXS) and other genetic diseases. The further understanding of the regulation of gene expression during neuronal development may provide new approaches for the diagnosis and treatment of these disorders.

Key words: neuronal development; gene expression; regulatory mechanisms; transcription; epigenetics; histone modification; DNA methylation; neurodevelopmental disorders; autism spectrum disorders

1 Introduction

Central nervous system development results from the interactions between intrinsic genes and extrinsic environment. The process of neuronal development consists of successive developmental stages including prolifera-
tion, differentiation, migration, axon extension, dendritogenesis and formation of functional synapses \([1, 2]\). In this developmental sequence, the interplay between genes and environment ensures that each step must occur in the proper timing and sequence. Their successful regulation requires that numerous groups of genes should be turned on and off in response to extracellular signals in an elegant spatiotemporal manner \([3]\). Thus, the regulation of gene expression plays a pivotal role in the neuronal developmental program. As the vertebrate brain continues its maturation long after birth, the protracted developmental process renders the vertebrate brain vulnerable to disruptions in this developmental program. Altered spatiotemporal expression of genes required for later neuronal development is likely to affect synaptic connectivity and organization of neuronal networks, leading to functional and behavior disorders \([2]\).

To better understand the spatiotemporal gene expression features in neuronal development, in the first part of this review, cerebellar granular neurons (CGNs) serve as an example to explain the various patterns of gene expression at different developmental stages. CGNs were chosen as a model system in the study of neuronal development because they possess the following unique advantages. First, CGNs are the most abundant neurons in the vertebrate brain, representing more than half of the total neurons \([4]\). Second, the differentiation process of CGNs has been well defined \([5]\). Third, CGNs play critical roles in cerebellum (Cb) functions via forming connections between mossy fibers, the major afferent input of the Cb, and Purkinje cells, the sole output of Cb \([5, 6]\). Moreover, besides motor coordination, increasing evidence indicates that the Cb also plays a significant role in cognitive functions, such as attention, language, emotional behavior, sleep and so on \([7, 8]\). Actually, this spatiotemporal pattern of gene expression in development is not limited to CGNs, and also can be found in other types of neurons \([9–12]\). In the second part, the regulatory mechanisms of gene expression in maturing neurons were reviewed, including the regulation by sequence-specific DNA-binding proteins (TFs and activators), epigenetic regulation (histone modifications, DNA methylation, chromatin remodeling), and the regulation by the newly emerged area of microRNAs (miRNAs). Localized mRNA translation is another regulatory mechanism in maturing neurons. It has been reviewed elsewhere \([13–15]\) and will not be further discussed here. The third part of this review is about neurodevelopmental disorders (NDs) that are related to the alterations of gene expression during neuronal development, including autism spectrum disorders (ASD), Rett syndrome (RTT), and Fragile-X syndrome (FXS). Advances of human genetics have revealed that some of these disorders are caused by mutations in transcription factors (TFs), transcriptional cofactors, epigenetic control proteins or miRNA \([16, 17]\). As the space limitation in the text, the detail physiological functions and anatomic structures of different neurons, and the treatment of related disorders are beyond this paper.

2 Spatiotemporal regulation of gene expression in neuronal development

During the neuronal development, the expression patterns of different groups of genes are precisely regulated in a spatiotemporal manner. For example, CGNs, which undergo a well-defined sequence of differentiation in Cb, show a fine spatiotemporal regulation of gene expression during development (Fig. 1A). In mouse, granule cell precursors (GCPs) originally arise from a dorsal hindbrain structure called the rhombic lip \([18]\). Before birth, GCPs leave the rhombic lip and stream across the outer surface of the Cb to form a region named external germinal layer (EGL). In the first two weeks after birth, cells in EGL undergo extensive proliferation to generate a large pool of GCPs \([19]\). Increasing numbers of GCPs gradually move inward, then exit the cell cycle and initiate differentiation by extending axons that form synapses with Purkinje cells \([20]\). With increasing age, the differentiating cells continue to migrate inward through molecular layer (ML) and Purkinje cell layer (PL) until they reach the final destination, internal granule layer (IGL). At 3 weeks of age, the EGL disappears and all GCPs complete their migration and differentiation into mature functional granule cells \([19, 21]\).

Associated with different morphological stages in CGN development, numerous sets of genes are sequentially expressed in different spatiotemporal manners \([3]\). Many of these genes are characteristic for a given stage of CGN maturation (Fig. 1A). For example, \(Math1\), which encodes a basic helix-loop-helix (bHLH) transcription factor, is expressed from embryonic day 9.5 in the upper rhombic lip and restricted to the EGL of
the Cb by E18. Math1 mRNA is undetectable in the postmitotic neurons in the following developmental stages [22]. The cell adhesion molecule Tag-1 is mainly expressed in postproliferative EGL cells and temporally down-regulated prior to onset of radial migration [23]. Gene Homer 2 is expressed in the migratory and immature IGL cells and slightly expressed in the mature cells [24]. However, most genes that encode proteins are involved in mature neuron functions, such as Gabra6, a subunit of gamma-aminobutyric acid (GABA) A receptor, is expressed in mature cells in IGL and is up-regulated during late development [25, 26].

Genome-wide gene expression analysis indicated that two striking groups of genes were temporally regulated during postnatal development in mouse CGNs (Fig. 1B) [27]. One group of genes are highly expressed at early postnatal stage and temporally down-regulated during development. The other group of genes are initially suppressed at early stage and then temporally up-regulated in the late maturing process [27]. Gene Ontology (GO) analysis indicated that the products of genes in the temporal-up group are significantly enriched in synapitic vesicle, dendritic spine, neuronal cell body, axon, synaptic membrane etc (Fig. 2A). Consistent to the component results, the function and process analysis indicated that the temporal-up genes are mainly involved in neuronal maturation and function, such as neuron differentiation, neuron development, neuron projection morphogenesis, neurotransmitter transport, cell adhesion and cell-cell signaling etc (Fig. 2B). On the other hand, the products of temporal-down genes are significantly enriched in nucleus, chromosome, condensed chromosome kinetochore, actin filament and cytoskeleton etc (Fig. 2C), participating processes of mitosis, cell cycle, cell division and migration, neurogenesis etc (Fig. 2D). (The link to the microarray data sets: http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?token=rtahnkckmemcarq&acc=GSE42018.) Diverse gene expression profiles in maturing neurons indicate that multiple gene regulation mechanisms occur in developmental program.

3 Regulatory mechanisms of gene expression in maturing neurons

Regulatory mechanisms of gene expression have been extensively investigated in neuronal development, including the regulation of TFs and co-activators, chromatin structure and posttranslational modifications. Non-coding RNA has been shown a new regulatory mechanism of gene expression in nervous system.
Fig. 2. Relative enrichment of temporally-regulated genes in cerebellar granular neurons during postnatal development. Temporally-regulated genes were identified by microarray analysis[27] and the data sets have been submitted to GEO data base: http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?token=rtahnkckmemcarq&acc=GSE42018. Gene Ontology (GO) analysis was generated by GENERIC GENE ONTOLOGY TERM FINDER at: http://go.princeton.edu/cgi-bin/GOTermFinder. On each panel, GO ID and term were labeled on left and the P values were labeled on right. A: The component of temporal-up genes. B: The function and process of temporal-up genes. C: The component of temporal-down genes. D: The function and process of temporal-down genes.
3.1 Transcriptional regulation by sequence-specific DNA-binding proteins

The first and the most rapid transcriptional changes are mediated by nuclear TFs or co-activators that sit pre-bound and primed at their target gene promoters or enhancers. In response to extrinsic stimuli, post-translational modifications occur to these TFs and then the changes of protein-protein interactions will rapidly recruit RNA polymerase and facilitate the transcriptional machinery assembling (Fig. 3A). For example, the transcription factor myocyte enhancer factor 2A (MEF2A) is highly expressed in CGNs throughout the period of synaptogenesis and it plays an important role in the synapse development, memory formation, and regulation the growth and pruning of neurons in response to stimulation. The sumoylation of MEF2A at lysine-403 forms a transcriptional repressor which promotes dendritic claw differentiation. On the other hand, MEF2A can be dephosphorylated at serine-408 by calcium signaling induced calcineurin (CaN) activity, thereby, promoted a switch from sumoylation to acetylation at lysine-403, resulting in a transactivation form and leading to inhibition of dendritic claw differentiation.

A second and slower transcriptional regulation is mediated by TFs that translocate to the nucleus following stimulation and result in the transcriptional machinery assembling. These TFs could undergo posttranslational modifications or nucleus translocation to target the promoters of late response genes, resulting in the second wave of gene expression in response to extrinsic stimuli.
Acta Physiologica Sinica, April 25, 2015, 67(2): 113–133

DNA binding domain and Ca²⁺ systems

ulation of genes central for many developmental

tor of activated T-cells (NFAT) family functions in reg

gates, nuclear factor I (NFI) plays a key role in orchestrating the neuronal activity-dependent gene regulation for terminal maturation of granule cells. The phosphorylation of ETV1 via the brain-derived neurotrophic factor (BDNF)-mediated ERK1/2 cascade up-regulates the battery of maturation genes, including those of the NMDA receptor subunit GRIN2C (glutamate receptor, ionotropic, NMDA2C) and TIAM1 (T-cell lymphoma invasion and metastasis 1) proteins.

3.2 Transcriptional regulation by epigenetics

In addition to the regulation of sequence-specific DNA-binding factors, research on neuronal transcriptional regulation in the past decade has been shifted to the investigation of mechanisms that regulate chromatin, including chromatin modifications, DNA methylation and chromatin remodeling. The regulation of mRNA abundance by miRNAs has also been extensively investigated in neuronal development. As these regulatory mechanisms that impact gene expression without changes of the DNA sequence, they are collectively named epigenetics.

3.2.1 Histone modifications

The physiological form of our genome is chromatin, which consists of genomic DNA and proteins. Nucleosome, the basic repeating unit of chromatin, is composed of approximately 147 base pair of superhelical DNA wrapped around a histone octamer consisting of two copies of each core histone H2A, H2B, H3 and H4. Amino (N-) and carboxy (C-) termini of histone proteins (histone tails) are subject to a different pattern of covalent modifications, including acetylation, methylation, phosphorylation, ADP-ribosylation, ubiquitination, etc. These different post-transcriptional modifications of individual histone protein occur at specific amino acid residues, such as acetylation and methylation at lysine (K) or arginine (R) residues and phosphorylation at serine (S) or threonine (T) residues, by the actions of certain enzymes (Fig. 4). Histone modifications represent a fine molecular code that modulates many cellular processes including neuronal development, plasticity and multiple forms of behavioral memory.

Acetylation is one of histone modifications that have been extensively investigated. Acetylation at several lysine residues throughout the N-terminal tails of core...
Histone proteins is generally associated with gene activation and is regarded as an epigenetic mark associated with dynamic chromatin [47]. Histone acetyltransferases (HATs) catalyze the addition of acetyl groups to histone lysine residues and histone deacetylases (HDACs) function to remove these modifications [54–56]. By neutralizing the positive charge of histone proteins, acetylation can effectively decrease the electrostatic affinity between histone tails and negatively charged DNA and loosen the chromatin structure. Finally, various effector proteins, transcriptional coactivators and members of the general transcriptional machinery are recruited and transcription is initiated [57–59].

Another important histone modification is methylation, which can exist in multiple valence states [e.g. mono- (me1), di- (me2), and tri-methylated (me3) forms] and exhibit slow turnover kinetics under normal cellular conditions [47]. The enzymes responsible for the addition and removal of methyl groups are histone methyltransferases (HMTs) and histone demethylases (HDMs), respectively [60, 61]. Histone methylation was once thought as a stable chromatin ‘mark’ that might act to control chromatin structure and the potentially related patterns of gene expression. However, as numerous sites and valence state-specific HDMs have been discovered, it is further complicated to categorize the methylation states as active or repressive for the gene expression [61]. For example, methylation of lysine 4 or 36 on histone H3 is related to transcriptional initiation and elongation, respectively, whereas methylation of lysine 9 or 27 on H3 is more strongly associated with transcriptional repression and silencing [47]. Moreover, methylation of lysine and arginine residues, some of which overlap with sites of histone acetylation, can be associated with either gene activation or repression depending on the residues being modified [62].

Fig. 4. Regulation of gene expression by epigenetics in neuronal development. Transcription activity can be regulated by histone modifications, such as acetylation (Ac), ubiquitination (Ub), phosphorylation (Ph), and methylation (Me), and each modification is catalyzed by one or more different enzymes. Nucleosome sliding driven by ATPase regulates the transcription activity by uncovering or masking the transcription factor binding sites. DNA methylation of CpG islands may result in stable silencing of gene expression via impairing transcription factor binding and/or recruiting repressive methyl-binding proteins. The target genes either encode neuronal proteins that function during neuronal development, or express transcription regulatory factors, which in turn regulate target genes expression at transcriptional level. Regulation at transcriptional level also regulates non-coding RNAs, such as miRNAs. By repressing translation or targeting mRNA degradation, miRNAs give rise to the fine-tune gene expression profiles. In the process of neuronal development, transcription factors, coactivators, suppressors, histone modifications, DNA methylations, and tons of miRNAs form a regulatory network on gene expression.
3.2.2 DNA methylation
Mammalia genomic DNA is also subject to methylation throughout the genome by the catalyzation of DNA methyltransferases (DNMTs) that transfer a methyl group from S-adenyl methionine (SAM) to the fifth carbon of a cytosine residue to form 5mC [63]. Three members of DNMTs have been identified, DNMT1, DNMT3a and DNMT3b. DNMT3a and DNMT3b can establish a new methylation pattern to unmodified DNA, while DNMT1 functions in DNA replication to copy the DNA methylation pattern from template DNA strand to the newly synthesized strand [64]. Most DNA methylation occurs on cytosines that precede a guanine nucleotide or CpG islands. The majority of gene promoters, approximately 70%, reside within CpG island [65]. Compared to other stretches DNA, CpG islands contain less nucleosomes and enhance the accessibility of DNA and promoter transcription factor binding [64, 66]. However, the methylation of CpG islands results in stable silencing of gene expression via impairing transcription factor binding and/or recruiting repressive methyl-binding proteins, such as MeCP2, which is mutated in the childhood neurological disorder RTT (see part 4.2 below) [67,68].

DNMT1 are expressed throughout neuronal development and plays a critical role in the regulation of gene expression in neuronal maturation [69]. For example, glial fibrillary acidic protein (GFAP), which is the main intermediate filament protein in mature astrocytes and also an important component of the cytoskeleton in astrocytes during development [70]. The DNA methylation of the Gfap promoter represses its expression in early neurogenesis at E11.5, coinciding with the high expression of DNMT1 [71]. At E14.5, the Gfap promoter underdoses DNA demethylation to coincide with the differentiation of the astrocytic lineage [72]. Between E8.5 and E13.5, a window period of neurogenesis, conditional knock out of DNMT1 causes hypomethylation of differentiating neurons and demethylation of the Gfap promoter in neural precursor cells, resulting in accelerating astrogliogenesis [73]. Hypomethylated neurons show multiple maturation defects including dendritic arborization and impaired neuronal excitability [74]. These studies demonstrated that precise regulation of DNA methylation is critical for differentiation and maturation of the central nervous system.

3.2.3 Chromatin remodeling
Another epigenetic control of gene expression is chromatin remodeling by ATP-dependent chromatin remodeling complexes, which use energy derived from ATP-hydrolysis to induce nucleosome sliding, resulting in uncovering or masking the transcription factor binding sites (Fig. 4) [75]. BRG and BRM, two of mammalian SWI/SNF2-like ATPases, are alternative subunits of chromatin remodeling complexes BAF (Brg/Brm associated factor) [76]. BRG and the npBAF subunits (BAF45a and BAF53a) have been shown to be required for neuronal stem cell/progenitor self-renewal and proliferation [77]. In addition, regulated expression of BRG/BRM has been implicated in orchestrating the developmental changes in gene expression program that underlie neurogenesis and dendrite outgrowth [78, 79]. Furthermore, BRG1 forms a complex with CREST (Ca2+-responsive transcriptional coactivator), which had previously been implicated in activity-dependent dendritic growth [80], plays a critical role in regulating promoter activation by orchestrating a calcium-dependent release of a repressor complex, and a recruitment of an activator complex [81].

3.2.4 Regulation of gene expression by miRNAs
miRNAs, the largest class of non-coding RNAs, are ~22 nucleotides (nts) molecules that act as post-transcriptional regulators of gene expression by repressing translation or targeting mRNA degradation [82]. The miRNAs are initially transcribed in nucleus by RNA polymerase II from intergenic regions, introns or exons of non-coding RNA genes, or within introns of protein-coding genes [83]. This primary (pri-) miRNAs can be recognized by nuclear protein Drosha, which cleaves the pri-miRNA from several kilobase (kb) into 70–80 nts with its RNase III activity under the direction of its co-factor DiGeorge syndrome critical region 8 (DGCR8) protein [84]. The processed precursor (pre)-miRNAs are then exported to cytoplasm, where Dicer further converts pre-miRNAs into the mature 22 nts RNA duplex [85]. The duplex is then unwound and one strand is preferentially loaded into the RNA induced silencing complex (RISC). By imprecise base pairing with sequences in the 3′ UTRs (three prime untranslated regions) of target mRNA, miRNAs modulate gene expression through transcript destabilization and translational attenuation [82,84]. Many miRNAs do not act as on-off switches, but rather fine-tune gene expression profiles [86,87]. The combined actions of miRNA and TFs are able to elegantly regulate gene expression on a global level in a manner that cannot be achieved by
TFs alone \[82\].

The accumulating evidence suggests that miRNAs are essential regulators of gene expression in neuronal development. In mouse, the depletion of miRNAs by inactivation of Dicer results in severe reduction of dendritic branch elaboration \[88\]. Numerous miRNAs have been shown to be presented in central nervous system, including miR-124 \[89\], miR-9 \[90\], miR-132 \[91\], miR-134 \[92\], miR-137 \[93\], miR-138 \[94\], miR-125b \[95\], miR-128 \[96\], miR-375 \[97\]. These neuronal miRNAs have been shown to play important roles at various stages in neuronal development and maturation, including neurogenesis, neurite outgrowth, dendritogenesis, and spine formation \[98\].

In the mouse brain, miR-124 is the most abundant miRNAs and it accounts for 25%-48% of all brain-expressing miRNAs \[99\]. In differentiating mouse P19 cells and mouse primary cortical neuron, overexpression of miR-124 promotes neurite outgrowth, while blocking miR-124 function delays neurite outgrowth \[100\]. By targeting Rho GTPase family members, cell division cycle 42 (Cdc42) and deactivating ras-related C3 botulinum toxin substrate 1 (Rac1), miR-124 diminishes F-actin density and stimulates tubulin acetylation and possibly regulates the cytoskeletal reorganization to promote neurite outgrowth \[98, 100\]. miR-124 is also an important regulator of the temporal progression of adult neurogenesis in mice subventricular zone (SVZ). By repression of SRY-box containing gene 9 (Sox9), miR-124 inhibits precursor amplification and stimulates differentiation \[89\]. Thus, miR-124 appears to play a major role in neuronal differentiation via down-regulating genes essential for precursor proliferation, whereas stimulating neuron-specific genes and cytoskeletal rearrangements \[98\].

The second example of a miRNA involved in neuronal development is miR-9, which is also highly expressed in the nervous system and highly conserved among species \[99\]. Overexpression of miR-9 in the developing brain has been shown to alter migration and differentiation of neuronal precursors, inducing premature differentiation \[90\]. By inhibiting taillless gene (TLX), a gene involved in the division of neuronal stem cells, miR-9 attenuated neuronal stem cell proliferation and enhanced differentiation. On the other hand, TLX inhibits miR-9-1, one isoform of miR-9, through binding to a 3’ genomic sequence to form a feedback loop regulation in neuronal differentiation \[90\].

In addition, by down-regulating of the forkhead box protein G1 (Foxg1), miR-9 promotes generation of Cajal-Retzius cells in the medial pallium of the developing telencephalon \[101\]. Besides Foxg1, several other TFs, such as Elav2, Gsh2, Nr2e1, and Pax6, have been shown to be down-regulated by miR-9 in the transcriptional regulation through various mechanisms \[102\].

Another two miRNAs, miR-132 and miR-134, that have been extensively investigated for their roles in neurite extension and synapse formation. The expression of miR-132 in rat hippocampus is temporally regulated in the early postnatal stage. Its level is low in the first postnatal week but accelerates between P7 and P21, coinciding in a period of active synaptogenesis \[98\]. On the contrary, the expression of P250GAP, an NMDA receptor-associated Rho GTPase-activating protein (GAP), which regulates spine morphogenesis, is high in early stage and decreases during maturation \[103\]. Interestingly, expression of miR-132 in neuronal cells is regulated by the cAMP response element-binding (CREB) protein, a key regulator in neuronal development, plasticity, and maturation. In response to extrinsic stimuli, induction of miR-132 by CREB promotes neurite outgrowth by targeting P250GAP \[104\]. Similarly, miR-134 is also regulated by activity at transcriptional level. The expression level of miR-134 in the rat hippocampus increases during development and reaches the peak at P13, when synaptic maturation occurs. By targeting Limk1, which regulates spine outgrowth by controlling cytoskeletal reorganization, miR-134 represses spine growth. However, this suppression can be alleviated by BDNF-induced synthesis of Limk1 \[48\].

In the process of neuronal development, these regulatory mechanisms on gene expression by epigenetics usually work together, or even combine the regulation of sequence-specific DNA-binding proteins (Fig. 4). For example, histone modification and DNA methylation affect the transcription of target neuronal genes or transcriptional regulatory proteins, such as TFs, coactivators, or suppressors. Subsequently, these factors could target other gene promoter regions, forming a regulatory network on transcription (Fig. 4). At posttranscriptional level, numerous groups of miRNAs regulate the abundance of mRNAs, the products of which can be either transcriptional regulatory factors or neuronal proteins that govern the neuronal development. Overall, the regulation of gene expression in neuronal
development is a complex network, which consists of TFs, coactivators, suppressors, histone modifications, DNA methylations, and tons of miRNAs. By targeting chromatin structure, transcription and translation processes, this network accomplishes the exquisite regulation of the neuronal gene expression in response to environmental stimuli at different developmental stages (Fig. 4).

4 NDs

Alterations of gene expression during neuronal development have been shown to be implicated in NDs, such as mental retardation, ASD, RTT, FXS, Angelman syndrome, Down syndrome, and other rare genetic diseases. These disorders can be caused by mutations in TFs, transcriptional cofactors or epigenetically control proteins. Here I took ASD, RTT, and FXS as examples to show the correlation between dysregulation of gene expression in maturing neurons and NDs.

4.1 ASD

ASD is classified as a pervasive developmental disorder (PPD) and characterized by some common clinical features, such as impaired social interactions, communication defects, and repetitive behaviors or a narrow range of interests [105]. In western countries, the prevalence of ASD has increased in recent surveys and current estimates of prevalence is around 20/10 000 [106], and the prevalence in children is around 1/100 to 1/150 [107]. In ASD prevalence, boys are typically affected at least four times more commonly than girls [105]. In 2014, the Centers for Disease Control and Prevention (CDC) released new data on the prevalence of autism in the United States. This surveillance study identified 1 in 68 children as having ASD, and 1 in 42 boys and 1 in 189 girls. In mainland China, as no national-wide systematic epidemiological studies have ever been reported, the number of people who were diagnosed with ASD remains unknown. Based on some local sample surveys in major cities, the estimated prevalence of ASD in children was about 11/10 000 to 16/10 000 [108, 109]. However, some researchers thought that this occurrence ratio of ASD in children in China was possibly underestimated because of differences in diagnostic criteria, finical constraints and technical limitations in rural areas [110, 111]. Thus, it is urgent to initiate a national-wide systematic epidemiological study on the prevalence of ASD in mainland China.

ASDs can be a major source of stress on the long-term health, and social and financial well-being of affected individuals, their families and society as a whole. Although the etiology and pathogenesis of ASD are not very clear yet, ASD is regarded as a highly genetic disorder [112]. It has been clearly demonstrated that ASD can be caused by many different genetic changes. More than 200 autism susceptibility genes, hundreds of copy number variations (CNVs), more than 100 chromosome fragile sites, and thousands of noncoding RNA molecules (snoRNA, miRNA, and piRNA) have been found to have a relationship to autism [113], indicating the etiology of ASD is complex and heterogeneous [114]. In addition, ASD has been noted as a comorbid feature of more than 100 genetic and genomic disorders, including RTT, Angelman syndrome, FXS, Down syndrome and other rare genetic diseases [115, 116].

Genetic studies demonstrated that up to 7%–10% of children with ASD have a variety of de novo chromosomal deletions and duplications [108, 117]. For example, X chromosome deletions implicated the NLGN3 and NLGN4 genes, which encode neuroligins 3 and 4 [118]. Neuroligins are synaptic adhesion molecules and regulate either the number or the function of synapses [119]. A 2-base-pair deletion in NLGN4 gene has been identified in a large family with many affected male patients showing mental retardation and/or autism [120]. Another strong genetic association for ASD is SHANK3 (SH3 and multiple ankyrin repeat domains 3), which encodes a scaffolding protein serving as a binding partner of neuroligins. A mutation of a single copy of SHANK3 on chromosome 22q13 can result in language and/or social communication disorders [121]. A genome-wide analysis of rare copy-number variation (CNV) in ASD families indicated that there could be 130–234 ASD-related CNV regions in the human genome, including rare de novo events at chromosomes 7q11.23, 15q11.2-13.1, 16p11.2, and Neurexin 1 [122]. By using whole-exome (coding regions of the genome) sequencing of substantial individuals, three recent studies identified roles of de novo exonic mutations in the pathogenesis in ASD, including SCN2A (sodium channel, voltage-gated, type II, alpha subunit) [123], KATNAL2 (katanin p60 subunit A-like 2), CHD8 (cadherin 8) [124], NTNG1 (netrin-G1), GRIN2B (glutamate receptor, ionotropic, N-methyl D-aspartate 2B), LAMC3 (laminin subunit gamma 3), and SCN1A (sodium channel, voltage-gated, type I, al-
pha subunit)\textsuperscript{[125]}. Alterations of trimethylated H3K4 (H3K4me3), a histone mark associated with transcriptional regulation, have been identified in another study\textsuperscript{[126]}. Compared to controls, excess spreading of H3K4me3 from the transcription start sites into downstream gene bodies and upstream promoters was observed specifically in neuronal chromatin from autism cases. Variable subsets of autism cases exhibit altered H3K4me3 peaks at numerous genes regulating neuronal connectivity, social behaviors, and cognition, often in conjunction with altered expression of the corresponding transcripts. These data indicate that loss or excess of H3K4me3 at hundreds of loci leading to dysregulation of gene expression implicated in ASD\textsuperscript{[126]}.

Studies of mutations in NLGN3/NLGN4 and SHANK3 genes provided the evidence that genes encoding proteins of the neuronal synapse are involved in ASD. In addition, synaptic activity-related signals which regulate protein synthesis within the synaptic spine have been shown to be involved in ASD\textsuperscript{[127]}. By mapping homozygosity in families with shared ancestry, three rare inherited autosomal recessive risk alleles for ASD have been identified, such as PCDH10 (protocadherin 10), DIA1 (deleted in autism 1) and NHE9 (Na\textsuperscript{+}/H\textsuperscript{+} exchanger 9)\textsuperscript{[128]}. Interestingly, PCDH10 and DIA1 are targets of transcription factor MEF2 (myocyte enhancer factor 2), which plays important roles in activity-dependent transcription in response to membrane depolarization\textsuperscript{[129]}. Gene NHE9 may be a target of transcriptional regulation by activity-inducible transcription factor Npas4 (neuronal PAS domain protein 4)\textsuperscript{[130]}. These data suggest that defective regulation of gene expression after neuronal activity may be a pathophysiological mechanism that is common to genetically heterogeneous causes of autism\textsuperscript{[128]}.

As mentioned above, ASD is a common comorbid feature of other genetic disorders, including some monogenic disorders. The studies of these monogenic disorders provided important clues about the mechanisms underlying ASD and also contributed to the understanding of the correlation between gene expression and developmental disorders. Currently, at least two different clusters of genes have been extensively studied for monogenic forms of ASD: 1) genes encoding proteins involved in regulating protein synthesis, or 2) genes encoding structural synaptic proteins\textsuperscript{[131, 132]}. In the following part, two monogenic NDs will be discussed.

### 4.2 RTT

RTT is a developmental disorder that predominantly affects females with prevalence around 1 in 10,000. Classic RTT is characterized by apparently normal early development for the first 6 to 18 months of age followed by developmental regression and loss of acquired skills, loss of speech, emergence of autistic features, loss of purposeful hand movements, microcephaly, seizure, ataxia, respiratory abnormalities, and mental retardation\textsuperscript{[133, 134]}. More than 95% of classic RTT cases are caused by mutations in human MECP2 gene, which encodes the X-linked methyl-CpG binding protein 2 (MeCP2)\textsuperscript{[135]}. By binding to methylated DNA, MeCP2 can read these epigenetic marks in genome and recruit transcriptional repressors and chromatin-modifying enzymes to silence gene expression\textsuperscript{[116, 137]}.

MeCP2 is ubiquitous in mammalian tissues and is highly expressed in the brain\textsuperscript{[138]}. Although neurons within the brain contain the highest levels of MeCP2, the expression in these cells is heterogeneous and tightly regulated according to the maturational stage of the individual neurons. The level of MeCP2 is up-regulated in a temporal manner in the course of postnatal neuronal development with the peak in a subpopulation of mature postmitotic neurons\textsuperscript{[139]}, suggesting MeCP2 is involved in the maturation of exiting neurons rather than the development of new neurons from precursor cells\textsuperscript{[140]}. Besides neurons in the brain, the expression of MeCP2 has been detected in glial cells\textsuperscript{[141]} and astrocytes\textsuperscript{[142]}, and has been shown to be critical for RTT pathogenesis by disrupting support for neuronal dendritic maturation.

MeCP2 consists of a methyl-CpG-binding domain, a transcription repression domain (TRD), two nuclear localization signals, and a C-terminal segment\textsuperscript{[133]}. The methyl-CpG-binding domain specifically binds to 5-methyl-cytosine throughout the genome. The TRD associates with histone deacetylase and transcription silencer corepressor Sin3A. The nuclear localization signals mediate the translocation of the MeCP2 into nucleus, and the C-terminal segment facilitates the binding of MeCP2 to the nucleosome core. These interactions cause histone deacetylation and chromatin condensation, resulting in the repression of transcription\textsuperscript{[133, 143]}. Interestingly, the expression level of MeCP2 in mature neurons is similar to that of core histones, suggesting MeCP2 may play a histone-like role in regulating chromatin and gene expression\textsuperscript{[144]}. Recent studies indicat-
ed that activity-dependent regulation of MeCP2 is also involved in the control of synapse development and behavior [116, 145]. Loss of function of MeCP2 in cells, especially in maturing postmitotic neurons, may cause inappropriate overexpression of genes that are not required, resulting in the potential damaging effect in the maturation of neurons. By targeted disruption of MECP2 gene in mice, many of neurodevelopmental phenotypes seen in human have been recapitulated [146, 147]. On the other hand, reintroduction of functional MeCP2 can rescue the developmental phenotypes resulted from MeCP2 deficiency [148], suggesting that the damage caused by a lack of MeCP2 in neurons is reversible. However, overexpression of MeCP2 by even 2-fold can cause a progressive neurological disease with cognitive impairment, speech and social problems [149], suggesting that MeCP2 levels must be tightly regulated in vivo and that too much as well as too little functional MeCP2 may cause NDs.

A new regulatory mechanism of gene expression by MeCP2 has been reported recently [150]. MeCP2 regulates gene expression at posttranscriptional level by suppressing nuclear microRNA processing. Specifically, MeCP2 binds directly to DGCR8, a critical component of the nuclear microRNA-processing machinery, and interferes with the assembly of Drosha and DGCR8 complex. Protein targets of MeCP2-suppressed miRNAs include CREB, LIMK1, and Pumilio2, which play critical roles in neural development [150]. These findings suggest that the interactions between various factors such as MeCP2 and miRNAs, form a complex network in the regulation of gene expression.

It is important to note that RTT phenotypes can be also caused by other genetic mutations besides MECP2, such as CDKL5 mutations in early onset seizure variant [151] and FOXG1 mutations in the congenital variant [152]. On the other hand, the same MECP2 mutation can cause significant variability of phenotypes in classic RTT and even other kind of NDs, such as autism [153, 154], Angelman syndrome (AS) [155], and X-linked mental retardation (XLMR) [156]. This indicates the multiple functions of MeCP2 in neuron maturation and the complex molecular mechanisms of etiology of NDs.

### 4.3 FXS

FXS is the most common known inherited cause of intellectual disability and ASD, and it typically results from transcriptional silencing of gene FMR1 (fragile X mental retardation gene) and loss of its product FMRP (fragile X mental retardation protein) [157]. Epidemiological studies estimated that the prevalence of FXS is 1/5 000 in males and approximately half as many as in females [158, 159]. Besides the moderate to severe intellectual disability and autistic features as mentioned above, FXS is characterized by seizures and/or epileptiform activity, hypersensitivity to sensory stimuli, attention deficit and hyperactivity, motor incoordination, growth abnormalities, sleep disturbances, craniofacial abnormalities, and macroorchidism (see, excellent review [159]). As FXS is the leading monogenetic cause of ASD, accounting for approximately 5% of ASD cases [160], it has been a valuable model for investigating the pathophysiology that may apply to ASD.

In the majority of FXS, the causative mutation is a trinucleotide (CGG) repeat expansion in FMR1 [161]. In human, the 5′-untranslated region of FMR1 is a polymorphic CGG repeats with the most common normal length of 30 repeats. However, in FXS patient, this repeat is expended more than 200 (referred to a full mutation), typically 800, leading to hypermethylation and epigenetic silencing of FMR1 and loss the encoded protein, FMRP. Alleles with an intermediate repeat length (60–200 repeats) are called premutations. Because the premutation alleles are unstable in meiosis, especially in female meiosis, the repeat length often increases from one generation to the next and has the chance to expand into the full mutation and leads to FXS [162–164].

FMRP is a selective mRNA-binding protein [165] and is highly expressed in neurons. It has been estimated that FMRP binds about 4% of total mRNA in the mammalian brain [160]. FMRP acts as a translational repressor to negatively regulate protein synthesis in the brain, especially at synapses in neurons [167]. Loss of FMRP leads to increased translation of FMRP-bound transcripts and impairs normal synaptic plasticity, which is believed to be the molecular basis of intellectual disability in FXS patients [163, 164]. Consistent with this hypothesis, several independent studies provided the view that many symptoms of FXS arise from a modest increase in synaptic protein synthesis [167–169]. Synaptosomes have increased both protein and mRNA levels of microtubule associated protein 1B (MAP1B), calcium/calmodulin-dependent kinase II alpha (α-CaM-KII), and activity-regulated cytoskeleton-associated protein (ARC), which play important roles in synaptogenesis and neuroplasticity [164, 167]. During neuronal
development, programmed FMRP expression represses the translation of MAP1B and is required for the accelerated decline of MAP1B level during active synaptogenesis. In Fmr1 KO neurons, the loss of FMRP results in misregulated MAP1B translation, leading to abnormally increased microtubule stability [168]. Interestingly, the excessive protein synthesis observed in Fmr1 knockout mice can be reversed by antagonists of mGluR5. Moreover, in these mice, both genetic and pharmacological inhibition of mGluR5 reverses synaptic and behavioral impairments, which are the basis of human clinical trials that aim to test the efficacy of mGluR5 antagonists for the treatment of FXS [159, 170].

There is significant overlap between FMRP targets and ASD candidate genes, such as NLGN3, NRXN1, SHANK3, PTEN, TSC1 and NFI [116, 132, 159], and most of them have been shown to participate in synapse plasticity and some of them have been discussed above.

Although mounting evidence suggests that FMRP functions to negatively regulate protein synthesis, the mechanisms of translational regulation by FMRP are not fully understood. One possible mechanism is that FMRP inhibits translation of its target mRNA by stalling ribosomes and through association with microRNA. Recent study found that FMRP interacts with the coding region of transcripts encoding pre- and postsynaptic proteins and reversibly stalls ribosomes specifically on its target mRNAs, suggesting FMRP dynamically represses translation in a complex consisting of target mRNAs and stalled ribosomes [171]. Moreover, given the majority of FMRP binding was found within coding sequence of mRNA, supporting the original hypothesis that FMRP represses translation by blocking elongation [172]. However, other studies indicated that FMRP may repress translation throughout many phases of translational regulation, including initiation stage [173, 174]. FMRP cooperates with translational machinery either to stall elongation or block initiation, in which the association with miRNA machinery has been shown to be involved. FMRP interacts with members of the RNA-induced silencing complex (RISC), including Argonaute and Dicer [175, 176]. Several specific miRNAs, such as miR-125b and miR132, are selectively associated with FMRP via interaction with members of the RISC [95, 177]. However, how ribosome stalling and miRNA-directed translational repression are spatially and temporally coordinated with each other remains to be investigated.

5 Concluding remarks

During neuronal development, gene expression can be regulated by multiple mechanisms at different levels of the flow of genetic information. By targeting chromatin structure, transcription and translation processes, the regulatory network consisting of transcription regulatory proteins, histone modifications, DNA methylations and miRNAs achieves the exquisite regulation of the neuronal gene expression in response to environmental stimuli at different developmental stages (Fig. 4). Dysregulation of the gene expression during neuronal development has been shown to be implicated in a number of NDs, suggesting the regulation of gene expression plays a crucial role in neuronal development and normal functions.

Currently, ASD is diagnosed entirely on behavioral criteria. To better target the underlying basis of ASD for diagnosis and treatment, considerable progress has been made in the identification of reliable biomarkers in genetics, neuroimaging, gene expression, and measures of the body’s metabolism [178–181]. However, as ASD encompass a range of neurodevelopmental conditions that are clinically and etiologically very heterogeneous, there are no reliable biomarkers with enough evidence to support routine clinical use unless medical illness is suspected. On the contrary, some available biomarkers for ASD are regarded doing more harm than good [182, 183]. Thus, we are still so far away to completely conquer these NDs like ASD, which relies on the breakthroughs in understanding the pathological mechanisms underlying these disorders and effective strategies for diagnosis and treatment.

The molecular mechanisms of gene expression during neuronal development will be continually focused on in biomedical research. Besides the classical molecular biological techniques that explore functions of gene/protein individually, approaches based on the next generation sequencing such as genome sequencing, RNA seq and ChIP seq make it possible to view the gene expression and regulation genome-wildly. These techniques greatly facilitate the understanding of the regulatory mechanisms in a global view and accelerate the identification of disease related mutations. The further understanding of the gene regulation in neuronal development under both physiological and pathological conditions may provide new approaches for diagnosis and pharmaceutical treatment of related disorders.
REFERENCES


acta physiologica sinica, April 25, 2015, 67(2): 113–133


85 Yi R, Qin Y, Macara IG, Cullen BR. Exportin-5 mediates the nuclear export of pre-microRNAs and short hairpin
110 Huang AX, Jia M, Wheeler JJ. Children with autism in the People’s Republic of China: diagnosis, legal Issues, and


DING Baojin: Gene Expression in Maturing Neurons


