# Axon Growth and Guidance in ASD: From Static Pathway Analysis to Dynamic Boolean Modeling

**Salvatore Contino**, Department of Engineering, University of Palermo, Palermo, Italy
**Giorgio Bertolazzi**, Department of Economics, Management and Statistics, University of Palermo, Palermo, Italy
**Francesco Calì**, Oasi Research Institute I.R.C.C.S., Troina, Italy
**Martina Cantone**, Laboratory of Systems Tumor Immunology, Department of Dermatology, University of Erlangen and Friedrich-Alexander University Erlangen-Nuremberg, Erlangen, Germany; Faculty of Mechanical Engineering, Specialty Division for Systems Biotechnology, Technical University of Munich, Munich, Germany
**Julio Vera**, Laboratory of Systems Tumor Immunology, Department of Dermatology, University of Erlangen and Friedrich-Alexander University Erlangen-Nuremberg, Erlangen, Germany
**Valentino Romano**, Department of Biological, Chemical and Pharmaceutical Sciences and Technologies (STEBICEF), University of Palermo, Palermo, Italy



**Background** 2
Systems Biology of Autism Spectrum Disorders and AGG: Review of Omic Studies 3
Dynamic Modeling of AGG: Continuous Versus Discrete Models 5
**Toward A Boolean Model of Growth Cone Dynamics** 6
In Silico Analysis of Growth Cone Dynamics: Model Building and Simulation 6
Simulations With Mutant Networks 8
**Discussion and Concluding Remarks** 10
**References** 12

## Glossary

**Autism spectrum disorder (ASDs)** Term that encompasses autism and similar disorders. More specifically, the following five disorders listed in DSM-IV: Autistic disorder, Asperger's disorder, PDD-NOS, childhood disintegrative disorder, and Rett's disorder.
**Autism** A disability; characterized by severe language and communication deficits, lack of normal relatedness, bizarre movement and self-stimulatory patterns, lack of normal handling of toys and other objects, and lack of most normal functional skills. Life Long developmental disability, neurological disorder affecting brain function.
**Axon guidance** The process by which growing nerve fibers find their targets in the developing brain. This process is regulated by a broad class of molecules that influence the direction of axon outgrowth through a variety of mechanisms.
**Basin of attraction** All states assumed by the nodes of a Boolean network that lead to a certain attractor form its basin of attraction.
**Bibliomics** A new concept representing a subset of high quality and rare information, retrieved and organized by systematic literature-searching tools from existing databases, and related to a subset of genes functioning together in "-omic" sciences.
**Bistability (multistability)** The ability of a model to rest in two (or more) different stable states (usually an "on" and "off" state).
**Chemotaxis (growth cone)** When a growth cone follows chemical signals (chemo-) to move toward (-taxis) a desired target.
**Comparative genomic hybridization (CGH) also array CGH** Technology wherein a DNA test sample is competitively hybridized with a reference sample of DNA of known sequence to a DNA microarray, used to detect copy number changes in the test sample.
**Copy number variants** Stretches of genomic sequence of roughly 1000 base pairs (1 kb) to 3 million base pairs (3 Mb) in size that are deleted or are duplicated in varying numbers.
**Emergent properties** Properties that arise from the interactions of many components and that were not apparent (that is, hard predict) when considering the individual properties of the components.
**Epigenomics** The study of how mechanisms other than changes in the underlying DNA sequence, such as the environment, can result in inherited changes in appearance (phenotype) or gene expression.
**Exome** The collection of exons, which are relatively small lengths of a whole genome and contain instructions for the body to build proteins—the building blocks that make up just about every type of tissue and fluid in the body. Exons account for about 1% of an entire genome but are thought to account for most inherited disorders.
**Feedback loop** When information from the end is also used to modify the process that produced it; in a feedback loop, information moves backwards to add into an earlier part of the pathway.
**Fragile X Syndrome** A genetic condition in which one part of the X-chromosome has a defect. The condition causes mental retardation. This disease is one of several monogenic diseases included in autism spectrum disorders.
**Gene ontology (GO)** A structured controlled vocabulary widely used for the annotation of gene products.

**Genome** The complete set of genes or genetic material present in a cell or organism.
**Genome-wide association study** A study that compares the whole genomes of study participants to find genetic variations associated with a particular disease.
**Genomics** The branch of molecular biology relating to the structure, function, evolution and mapping of genomes.
**Growth cone** Growth cones are located at the tip of a growing and navigating axons that senses and uses chemical signals to find its targets. They are highly dynamic, with growth and reorganization of actin filaments and microtubules constantly reshaping their structure on timescales of less than a minute. The growth cone consists of central and peripheral domains. The former contains microtubules extending from the axon shaft and a network of actin. The peripheral domain contains actin filaments and is characterized by finger-like filopodia joined by veil-like lamellipodia.
**Intellectual disability/mental retardation** A disability characterized by significant limitations both in intellectual functioning and in adaptive behavior as expressed in conceptual, social, and practical adaptive skills.
**Mendelian disorder (single-gene disorder)** A trait or disease that follows the patterns of inheritance that suggest the trait or disease is determined by a gene at a single locus.
**Methylation** Covalent attachment of methyl groups to DNA, usually at cytosine bases. Methylation can reduce transcription from a gene and is a mechanism in X-chromosome inactivation and imprinting.
**Microarray** A technology used to study many genes simultaneously, usually consisting of an ordered microscopic pattern of known nucleic acid sequences on a glass slide. In a common type of microarray, a sample of DNA or RNA is added to the slide and sequence-dependent binding is measured using sensitive fluorescent detection methods.
**Neurodevelopmental disorders** Neurodevelopmental disorders are a group of heterogeneous conditions characterized by delay or disturbance in the acquisition of skills in a variety of developmental domains, including motor, social, language, and cognition.
**Next generation sequencing (NGS)** Massively parallel high-output sequencing of nucleic acids, enabling cost-effective analysis of large numbers of target sequences in a single experiment.
**-ome, -omics** Suffixes indicating the complete data (and the study of that data) of a particular type (indicated by the root) from particular biological systems (e.g., genome/genomics, proteome/proteomics).
**Ontology** A controlled vocabulary that defines terms and their relationships for a specific knowledge domain. Although ontology and data model can be used interchangeably, an ontology emphasizes the vocabulary used for the entities and their relationships, whereas a data model emphasizes its structure. In computer science, a model of a particular domain, amounting to a logical description of a set of related concepts.
**Pathfinding** The process of the axons finding the right neuron or target to connect to by the formation of synapses (Neuropathfinding).
**Signal transduction (ST) and signal transduction pathways (STP)** ST is a communication process within a cell to coordinate its responses to an environmental change. The response is a reaction of the cell, e.g., the activation of a gene or the production of energy. A STP is a directed network of chemical reactions in a cell from a stimulus (an external molecule which binds to a receptor on the cell membrane) to the response (e.g., a gene whose activity is changed due to the binding of external molecule). The signal transduction network of a cell is the complete network of all signal transduction pathways.
**Whole-exome sequencing** A laboratory process that determines, all at once, the entire unique sequence of an organism's exome. Sometimes also called "exome capture."
**Whole-genome sequencing** A laboratory process that determines, all at once, the entire unique DNA sequence of an organism's genome.

## Background

Axon growth and guidance (AGG) are two tightly regulated and related processes required for the correct patterning of neuronal connectivity during embryonic development and repairing of nervous tissue following axonal injury (Kolodkin and Tessier-Lavigne, 2011; Stoeckli, 2018; Tessier-Lavigne, 2002; Tessier-Lavigne and Goodman, 1996). A key player in AGG is the growth cone (GC), a very dynamic structure located at the tip of a growing and navigating axon (Dent et al., 2011; Lowery and Van Vactor, 2009).

In spite of the intense research activity and important achievements of the past decades, the underlying cytoskeletal machinery of growth and guidance are not fully understood. For this reason, it is advisable to approach this biological event from a new perspective to unveil new regulatory mechanisms. Given the intricate regulatory network, the innovative method proposed by systems biology (SB) (Kitano et al., 2005) might comprehensively analyze AGG and GC dynamics by systemically inspecting the underlying biochemical networks as interdependent and connected systems. Systems Biology is an emergent paradigm in molecular and cell biology, in which the properties of the gene and proteins networks underlying key cell phenotypes are investigated by integrating and analyzing quantitative high throughput data with computational models. Under this premise, the corresponding morphological and dynamical phenotypes appearing during neuronal differentiation can be viewed as emergent properties of the AGG-GC system. Subsequent analysis procedures can associate such properties to the overall behavior of the underlying interconnected network of molecules, rather than the function of individual genes or proteins (Cohen and Harel, 2007).

When studying AGG from a systemic perspective it may be useful to distinguish between two levels of complexity. A first level concerns the expression of several cellular or sub-cellular phenotypes displayed in axon pathfinding some of which include: (i) GC steering, retraction and collapse, (ii) GC filopodia and lamellipodia formation and disappearance, (iii) axon growth, fasciculation and branching, (iv) neurite outgrowth during the first stages of neuroblast differentiation and (v) symmetry-breaking in neuroblasts leading to the formation of one axon and many dendrites on the two opposite sides of the cell. A second level of complexity concerns the molecular events occurring inside and outside the growing and navigating axon that are correlated to the phenotypes mentioned above. To date, a large number of molecules have been implicated: on May 2019 the Gene Ontology database counted at least 355 human proteins directly or indirectly involved in regulation of AGG. The annotated functions can be roughly grouped in three categories: (i) extracellular cues and relative specific membrane receptors, (ii) signal transduction (e.g., GTPases with their GEFs and GAPs, kinases, phosphatases), and (iii) effectors of cytoskeleton dynamics such as actin, tubulins and their direct interactors.

Although these three categories are common among different neuronal types in various species, they cover a central role for they have been proven to promote repulsive, attractive and growth responses of the GC: in fact, different signaling cascades converge to the three Rho-GTPases (Cdc42, RhoA, Rac) or to their GEFs and GAPs (Hall and Lalli, 2010). From the three Rho-GTPases, the transduced signals are conveyed to the cytoskeletal tubulins and actins through regulators promoting or inhibiting the formation of GC filopodia and lamellipodia (see Fig. 1). Furthermore, the role of signaling mediator of the Rho-class small GTPases is exploited in many basic cellular processes of brain development, from neurogenesis or neuronal migration to synaptogenesis and synaptic plasticity, beside axon guidance, thus indicating that altering this signaling cascade can potentially disrupt essential events in neurodevelopment. Other molecules such as cAMP, cGMP, and microRNAs as well as calcium do play important roles in AGG (Forbes et al., 2012; Gasperini et al., 2017; Iyer et al., 2014; Piper et al., 2007).

In humans, several genetic disorders induce dysfunctions of the protein network underlying AGG (e.g., Engle, 2010; McFadden and Minshew, 2013; Sbacchi et al., 2010; Wang et al., 2018). Relevant examples include corpus callosum agenesis, L1 syndrome, Joubert syndrome and related disorders, and several others reviewed in Engle 2010 (Engle, 2010). Since the publication of Engle many other studies reported that a dysfunctional AGG underlies several other neurological and psychiatric disorders including intellectual disability (Zamboni et al., 2018), autism spectrum disorders (ASDs) (McFadden and Minshew, 2013), schizophrenia (Wang et al., 2018), and epilepsy (Koyama and Ikegaya, 2018). Genes found mutated in these disorders can encode axon GC ligands and their receptors, downstream signaling molecules, axon transport motors and GC cytoskeleton proteins. For example, dysregulation of RhoA, Rac1/Rac3, and Cdc42 Rho-GTPases activity has been related to intellectual disability and other neurodevelopmental disorders including ASDs (Zamboni et al., 2018). Among the above mentioned disorders, ASDs is the neuropsychiatric disease where the link with AGG genes has been repeatedly demonstrated and for this reason we will specifically refer to it as an especially instructive case study.

## Systems Biology of Autism Spectrum Disorders and AGG: Review of Omic Studies

Structural and functional genomic studies performed in recent years have provided evidence for dysregulation of AGG in the etiology of ASDs. In this paragraph we briefly review few omic-based studies that have supported the association of AGG's perturbations with ASDs.

*Copy number variants (CNVs)*. Sbacchi et al. (2010) identified genes presenting CNVs through four large ASDs microarray data set and then used Gene Ontology and pathway analyses to determine common functions of duplicated or deleted genes only. They identified a substantial number of genes related to axonal guidance as well as morphogens (BMP, Wnt, and Engrailed) which are also known to participate in axon guidance (Charron, 2005) and to be linked to ASDs by previous studies (Kalkman, 2012). Gilman et al. (2011) analyzed rare de novo CNVs associated with autism and discovered a functionally connected gene-network related to synaptogenesis, axon guidance, and neuronal motility. Genome-wide association studies (GWAS). Hussman et al. (2011), similarly identified a substantial group of genes (potentially related to ASDs) involved in neurite outgrowth by applying GWAS *Transcriptomic and epigenetic studies*. Decreased expression of axon-guidance receptors has been detected in the anterior cingulate cortex in post-mortem brains of autistic patients (Suda et al., 2011). Tangsuwansri et al. (2018) investigated the epigenetic regulatory networks associated with ASDs by integrating global Long-Interspersed-Nuclear-Elements (LINE-1) methylation with gene expression profiling analyses. They found that LINE-1 insertions were significantly associated with differentially expressed genes (DEGs) in ASDs, and pathway analysis of LINE-1-inserted DEGs were associated with axon guidance signaling. Moreover, they observed that the LINE-1 methylation level was significantly reduced in lymphoblastoid cell lines from ASDs individuals with severe language impairment and was inversely correlated with the transcript level of LINE-1. The methylation level of LINE-1 was also correlated with the expression of the LINE-1-inserted DEG C1orf27 *Exome sequencing*. Griswold et al. (2015) performed targeted massively parallel sequencing of genes associated to ASDs in a case-control designed study and identified rare variant of genes involved in excitatory neurotransmission and neurite outgrowth and guidance pathways. Codina-Solà et al. (2015) identified rare inherited mutations enriched by integrating whole-exome sequencing (WES) and blood cell transcriptome RNA-Seq data in a subset of male patients with idiopathic ASDs, and these mutations have been later related to PI3K-Akt signaling and the axon guidance pathways.

This brief survey of the recent literature shows that defects of axon guidance genes are found in ASDs patients, and their involvement has been supported by the use of several high throughput (HT) techniques continuously improved especially in the last decade.

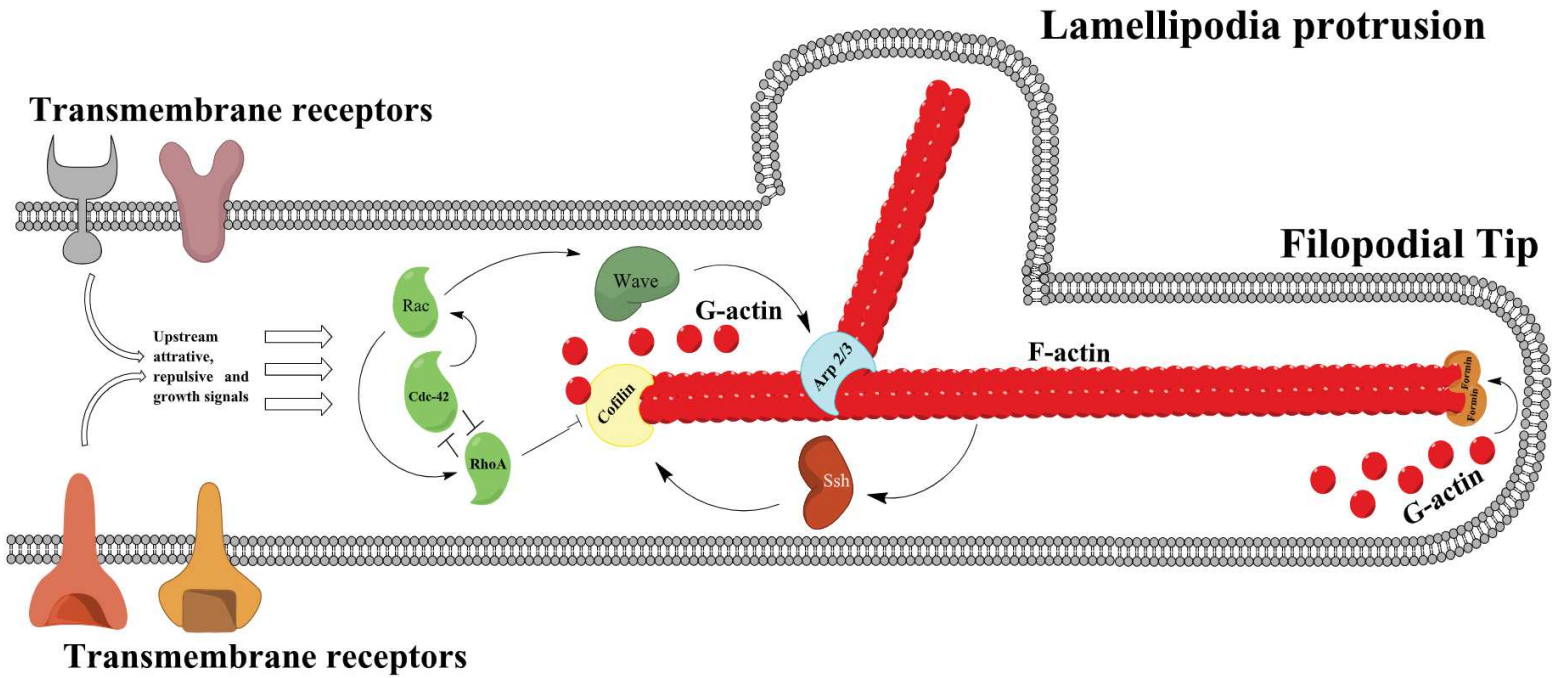



**Fig. 1** Simplified model of growth cone (GC) filopodium and lamellipodium showing a subset of the proteins regulating actin polymerization and depolymerization. Most of the extracellular signals interacting with membrane receptors induce signaling cascades that converge to the three Rho-GTPases to regulate GC dynamics. The node sequence F_actin > Slingshot (SSH) > Cofilin > G_actin > F_actin represents one of several feed-back loops (FL) present in the protein network. In particular, the above sequence represents a positive FL motif involved in the formation of cyclic attractors in the Boolean model simulation (see also Fig. 4).

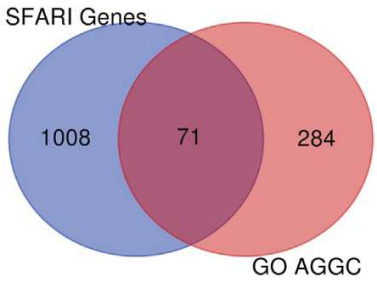



**Fig. 2** Venn diagram showing the intersection of 71 genes between SFARI genes conferring susceptibility to ASDs and the human genes associated to the "axon guidance" and "growth cone" GO terms (GO AGGC).

In order to better characterize the overall impact of axon guidance genes on autism, we here resort to an important web resource implemented by the Simons Foundation Autism Research Initiative: (SFARI). Within the SFARI framework, "SFARI Gene" (Abrahams et al., 2013; Banerjee-Basu and Packer, 2010) is a continuously updated database dedicated to the autism research community and comprises genes implicated in autism susceptibility. Referring to the state of the database in May 2019, the database archives 1079 genes, and each of them is described with a score reflecting the strength of the evidence linking it to the development of autism. Six categories have been defined: (1) High confidence; (2) Strong candidate; (3) Suggested evidence; (4) Minimal evidence; (5) Hypothesized but not proven; (6) Evidence does not suggest a relation. An additional category (S) holds genes constituting a risk factor for autism in the context of a syndromic disorder (e.g., fragile X syndrome). On the top of this, if a gene is classified as "S" and evidences relate it to idiopathic autism, the scoring system is again applied to measure the strength of that evidence (e.g., 3S, which would be listed both in the "S" category and in category 3). We focused on the genes classified "S" and searched for those genes associated to the "axon guidance" and "growth cone" GO terms (355 human genes on May 2019). As the Venn diagram in Fig. 2 shows, the intersection between SFARI and these GO terms lists consists in 71 genes.

Confirming the actual involvement of these mutations in the pathogenesis of ASDs is a challenging task since (i) these mutations are often very rare, i.e., detected in few patients thus jeopardizing the stronghold of reproducibility of scientific results and (ii), it is difficult to evaluate the functional impact of missense mutations on protein function. On the one hand, in silico Bioinformatic tools such as wAnnovar or Polyphen-2 represent a first approach in selecting potential pathogenic candidates based on in silico predictions of the effect of missense mutations on protein structure. On the other hand, though the use of in vitro or in vivo models would represent an important alternative approach to functionally test these mutations, they are still employed in a limited number of cases because the use of these models is technically demanding. Not to mention the additional difficulties encountered to assess in vivo and/or in vitro, the effects of a specific mutation in a given protein on other proteins belonging to the same functional network. Moreover, this complexity scale-up when considering that ASDs might be related to simultaneous mutations in different genes. In line with this, understanding the genotype-phenotype relationship in ASDs must consider the epistatic interactions between genes.

Together with protein folding simulations, Systems Biology approach suggests additional in silico approaches to extract biologically relevant information from a massive amount of data, especially from HT data. These approaches, based on bottom-up analyses and dynamic simulation, have been successfully used already in other biological and biomedical systems (e.g., in cancer Fumiã and Martins, 2013), but—to our knowledge—they have never been applied to autism and other neurodevelopmental disorders. In the remaining part of this article we will describe in more details these methods and present an example of how they can be applied to ASDs using AGG as a magnifying lens to zoom in one of several possible pathogenic mechanisms.

## Dynamic Modeling of AGG: Continuous Versus Discrete Models

Given the complexity of the AGG and GC biochemical networks, it becomes necessary to turn to Systems Biology (SB) methodologies since the simultaneous study of proteins dynamics is a demanding task if performed with available cytological and molecular technologies. In a nutshell, in Systems Biology knowledge about the systems investigated concerning the key genes, proteins and interactions is collected from databases and manual curation of scientific literature, but also making use of bioinformatics algorithms. With this information one derives a regulatory map, a graphical depiction of the network investigated where nodes are genes/proteins and edges are interactions. Making use of some heuristics, this regulatory map is translated into a computational or mathematical model. The basic feature of the computational model is that it can be used to make computational simulations of the biochemical network considered in different biomedical contexts. These simulations can be used to get new insights into the regulation of the system, to derive hypothesis or design experiments, but also to look for biomarkers and drug targets. This approach has achieved remarkable success specially in molecular oncology. Two additional ideas are that (a) the model can be better characterized (a.k.a. calibrated or trained) by matching model simulations with existing data before making actual predictive simulations and (b) there are different frameworks for computational modeling (i.e., ODEs, Boolean, agent-based, etc.), which differ in their trade-off in terms of biological detailness, scalability to large networks or accuracy of the simulations.

At the core of SB, mathematical and computational modeling cover an important role to reproduce in silico and tackle the biological variations of all molecules of a defined network (Fischer, 2008; Ingalls, 2013; Vallabhajosyula and Raval, 2010). Since in vitro/in vivo studies on AGG constitute burdensome and arduous tasks, mathematical simulations are especially suited because of their predictive power of the dynamics of biochemical systems with many molecules. In general, the dynamics of a network can be mathematically modeled either by continuous or discrete representation: in continuous models the variation of the concentration of each component (proteins, RNAs, etc.) is described with differential equations, while in discrete models, each element is characterized by discrete states. Referring to Boolean models, there are only two mutually exclusive states, the logic 1 and 0. To our knowledge, mathematical and computational models so far used to study AGG at both cellular and molecular levels, belong to the category of continuous model (e.g., Aletti and Causin, 2008; Krottje and van Ooyen, 2007; Padmanabhan and Goodhill, 2018; Roccasalvo et al., 2015; Verleysen, 2013). However, these studies have been employed to reproduce the simultaneous behavior of a limited number of proteins because of the necessity of determining kinetic parameters before running simulations. Despite the simplification offered by Boolean models and their extended usage to analyze the dynamics of various biological networks in different species (e.g., Albert and Othmer, 2003; Albert and Robeva, 2015; Hetmanski et al., 2018; Huang, 1999; Irurzun-Arana et al., 2017; Li et al., 2004), no studies have applied this type of modeling to study AGG.

## Toward A Boolean Model of Growth Cone Dynamics

### In Silico Analysis of Growth Cone Dynamics: Model Building and Simulation

The steps required to construct a Boolean model of the growth cone and to perform a dynamic simulation are shown in the flowchart of Fig. 3 and in the Box 1.

As a first step an extensive literature search is performed for identifying proteins and other molecular species proven to drive GC dynamics. This activity has been recently defined as part of a omic disciplines and named "Bibliomics." Specifically, we looked for interaction with attractive, repulsive or growth extracellular molecular cues. Overall, more than 80% (22/27) of protein data collected from the literature and databases surveys concerned rodent cortical neurons. Using the "CellDesigner" software we have then constructed several graphical maps (not shown) representing the "classical" signaling cascades involved in axon guidance (i.e., Netrin/DCC, Semaphorin/Plexin, Slit/Robo, Ephrin/Eph). Since all these pathways converge on the three Rho-GTPases (RhoA, Cdc42, Rac), we focused on the downstream part of the network and considered the three Rho-class small GTPases as the input variables for the simulations. The reduced network that we referred to as Default Network (DN) consists of 27 proteins which include the three Rho-GTPases and 24 more proteins directly involved in the regulation of the actin and tubulin cytoskeleton of the GC. Two so-called phenotypic nodes (FE = filopodia elongation and Lam = lamellipodia) were also added to facilitate the functional interpretation of attractors in terms of GC behavior.

The chemical reactions were then translated into Boolean functions according to the syntax of the "BoolNet" R package. The obtained Boolean model was used to study the attractors of the system. We simulated all possible initial states of the network (229) and obtained 72 attractors using a synchronous simulation. Over the 72 attractors, 40 were classified as fixed-point attractors and 32 as cyclic attractors. The latter ones can be further distinguished in bi-stable states (25%) and tetra-stable states (75%). In these

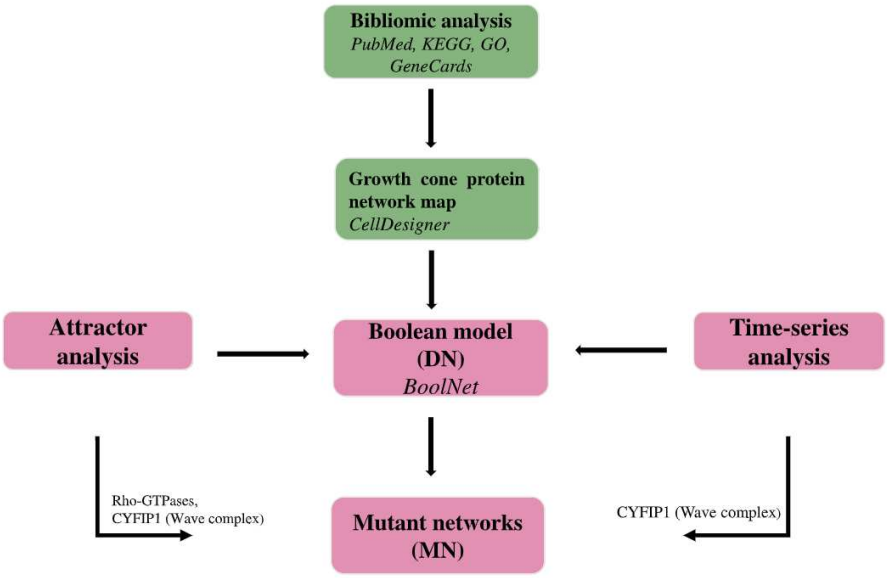



**Fig. 3** Flow-chart of the study design and its milestones. *Green-shaded* boxes refer to network construction, whereas *magenta-shaded* boxes refer to network analysis. *DN*, Default network.

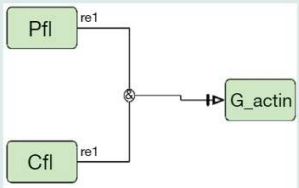
## Box 1 Boolean networks: a primer

Boolean networks and Boolean variables. Boolean networks models have been applied since 1969 (Kawano et al., 2005; Kitano et al., 2005) to represent gene networks and other types of biomolecular networks to study complex and non-linear regulatory dynamics. A Boolean model is a collection of Boolean variables that describe the state of the nodes (genes or proteins) at specific time points. Each Boolean variable is represented by n Boolean values $X_i$

$$X = \{X_1, X_2, \ldots, X_n\}$$

where $X_i$ can assume either the value 1 or 0 corresponding to its active or inactive state respectively.

*Transition functions.* Boolean variables are associated to network components and transition functions (or "rules"). Transition function are the Boolean analogous of molecular interactions or regulatory processes. A transition function is written with the three logical operators AND, OR, NOT which are comparable to product, sum, and negation operations respectively. For example, if A and B are necessary to activate C the corresponding Boolean rule is: C = A AND B. If only A or B are each sufficient to activate C the corresponding rule would be: C = A OR B. Finally, if A inhibits C whereas B activates C the corresponding rule would be C = B AND NOT A. In our Boolean network an example of transition function involving an "AND" is the following: "G_actin* = Profilin (Pfl) AND Cofilin (Cfl)." The corresponding graphical symbols and truth table for this function are shown in this Box.




| G_actin* = Pfl AND Cfl | | |
|---|---|---|
| Pfl | Cfl | G_actin* |
| 0 | 0 | 0 |
| 0 | 1 | 0 |
| 1 | 0 | 0 |
| 1 | 1 | 1 |

*Model simulations and network propagation.* In Boolean networks we can distinguish three types of nodes: (a) inputs: nodes that affect other nodes but are not affected by any part of the network, in a biological sense they can be signaling cues or proteins external to the network modeled; (b) output, nodes whose activation status is affected by the network but do not affect the network, in a biological sense they can account for phenotypes triggered by the network; and (c) nodes belonging to the intermediate layers: Nodes within the circuitry of the regulatory network whose activation status is affected by the network and also influence the status of other nodes. The state of a given node (1 or 0) is updated at each step of the simulation and it depends on the defined transition function.

*Expert knowledge-based construction and curation of Boolean networks.* A major advantage of Boolean networks is the fact that natural-language statements can easily be transferred into this representation. This allows for assembling Boolean networks entirely from expert or prior knowledge, i.e., the collection of statements on protein dependencies from literature and other sources, and expressing them as Boolean rules.

*Attractors in Boolean representations of biochemical networks.* In biochemical networks, Boolean variables represent the nodes (genes, proteins, etc.) and they can assume either the logic values 1 or 0 which are interpreted as two mutual exclusive biological conditions (present/absent, active/inactive, etc.). Although this might be considered as oversimplification, BNs have been extensively applied for their capability of approximating the real nature of different types of biological networks (Charron, 2005; Huang, 1999) and mathematical simulations of BNs can unveil new dynamics of the system under analysis. For example, the establishment of a state of the network or the cyclic repetition of particular states (technically indicated as "attractors") can be associated to specific behaviors or phenotypes (Kitano et al., 2005). Steady-state attractors are attractors that consist of only one state. Cyclic attractors consist of a set of states whose synchronous transitions form a cycle. The detection of attractors can be done with two different updating methods of the transition rules named synchronous and asynchronous.

*Synchronous versus asynchronous update.* With synchronous method, the transition function is applied on all at the same time (hence synchronously) while, for asynchronous method, at each time point of the simulation one transition function is selected randomly and therefore, only one element of the model is updated. This latter method aims to simulate the biological variance: each biochemical reaction is characterized by its speed rate, hence expression levels or activity of proteins are likely not to occur simultaneously. With synchronous simulation only two types of attractors can be reached: fixed-point (or steady state) and cyclic.

*Network simulation initialization.* For what concerns the initial condition of the simulation, a specific set of variables states can be selected, otherwise simulations are randomly initialized to test as many combinations as possible. This latter approach can be chosen when not enough biological information is available to characterize initial condition of the AGG/GC system. By systematically simulating different initial conditions, it has been shown that the same attractor can be reached from distinct starting points thus generating the basin of attraction. Biological relevant phenotypes are often associated to vast basins of attraction (Hussman et al., 2011; McFadden and Minshew, 2013).

*Software tools for Boolean model construction and simulation.* Boolean rules can be adapted to the syntax of a suitable software such as BoolNet (Suda et al., 2011) which runs on a R environment (Poole et al., 2016) to implement the dynamic analyses. A detailed description of the functions available in this can be found in the BoolNet package vignette and Reference Manual. For example, one function used to perform attractor analysis is getAttractors().

*In silico mutagenesis.* This procedure is aimed to reproduce the biological knock-out (KO) or over-expression (OE) (i.e., perpetual activation) of one or more genes. In the KO setting, the state of the selected node is fixed to 0 for the whole simulation, whereas OE is implemented by fixing the state of the node to 1. generateTimeSeries() is another function which allows to generate a large number of random initial states from which simulations are run for a user-defined number of time steps.

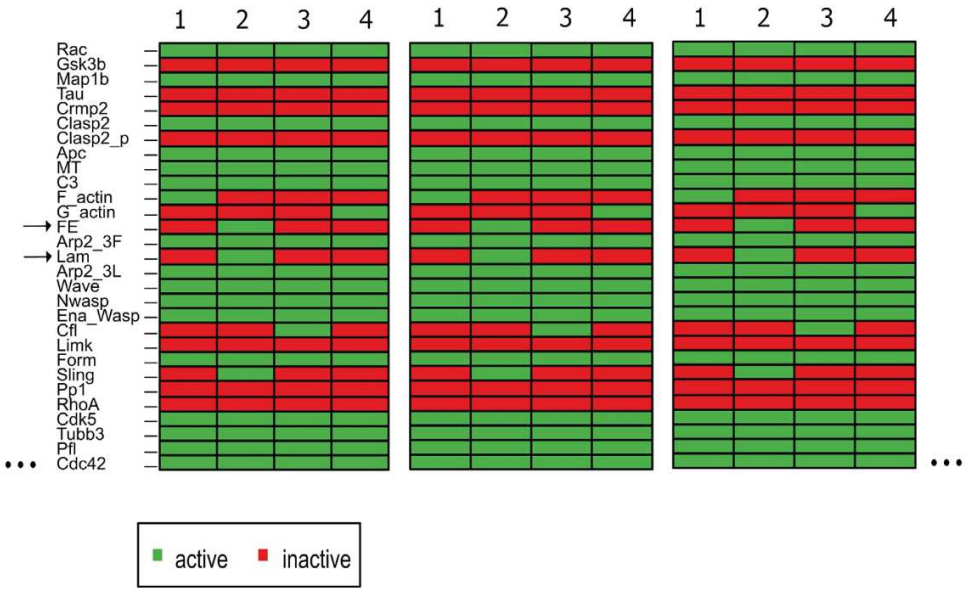



**Fig. 4** A cyclic attractor with four steps (1, 2, 3, 4) obtained by performing a synchronous exhaustive simulation of the default network (DN) of the growth cone containing 29 nodes (27 proteins + 2 phenotypic nodes). *Arrows point* to the two phenotypic nodes: lamellipodia (Lam) and filopodia (FE). Protein acronyms' key: *Rac*, Ras-related C3 botulinum toxin substrate 1 (Rho family, small GTP-binding protein Rac1); *Gsk3β*, glycogen synthase kinase 3-beta; *Map1b*, microtubule-associated protein 1B; *Tau*, microtubule-associated protein tau; *Crmp2*, dihydropyrimidinase-like 2; *Clasp2*, cytoplasmic linker-associated protein 2; *Apc*, APC regulator of WNT signaling pathway; *MT*, microtubules; *C3*, complex formed by: Tubb3 (tubulin beta 3 class III) and Map1b; *F-actin*, filamentous actin; *G-actin*, globular actin (monomer); *FE*, filopodia elongation; *Arp2/3*, actin-related protein complex; *Lam*, lamellipodia; *Wave*, WASP family verprolin-homologous protein; *Nwasp*, neural Wiskott–Aldrich syndrome protein; *Ena/Vasp*, Ena (Drosophila enabled)/VASP (vasodilator-stimulated phosphoprotein); *Cfl*, cofilin; *Limk*, LIM domain kinase; *Form*, formin; *Sling*, SlingShot protein phosphatase; *Pp1*, protein phosphatase 1; *RhoA*, Ras homolog family member A; *Cdk5*, cyclin-dependent kinase 5; *Tubb3*, tubulin beta 3 class III; *Pfl*, profilin; *Cdc42*, cell division cycle 42.

cyclic attractors, Lam and FE nodes display an "oscillating behavior" between 1 and 0 states. This behavior was observed in 50% of cases for FE and 100% of cases for Lam referring to cyclic attractors with 4 states (see Fig. 4 for an example of cyclic attractor generated by "BoolNet").

Among the 40 attractors classified as "fixed-point," 8 (20%) had at least one active phenotypic node (FE or Lam), 4 (10%) had both active phenotypic nodes (FE and Lam) and the remaining 32 (80%) attractors had both inactive phenotypic nodes. Recalling the theory of Kauffman (1993) in which multi-state attractors can correspond to a series of cellular phenotypes we hypothesize that (i) the oscillating behavior displayed by Lam and FE nodes in cyclic attractors evokes the appearance/disappearance behavior of the GC filopodia and lamellipodia in vivo and in vitro (Henle et al., 2011), (ii) the fixed-point attractors presenting both inactive FE and Lam nodes recall GC retraction or collapse, and (iii) cyclic attractors and fixed-point attractors presenting both Lam and FE active may correspond to GC constantly advancing or changing direction (steering) toward an attractive extracellular cue. Indeed, filopodia and lamellipodia appearance/disappearance in GC is caused by polymerization/depolymerization of F-actin and/or microtubules.

## Simulations With Mutant Networks

To show how the Boolean model of the GC can be used to check the effects of mutations in silico, we studied the attractors and the sequence of the Boolean network's states in relation to mutation introduced in the network. More precisely, we simulated the effects of knock-out (KO, Boolean state 0) and over expression (OE, Boolean state 1) of the three Rho-GTPases (i.e., Rac, Cdc42, RhoA) and the knock-out of CYFIP1 (Wave complex) genes (in this precise context, simulations of over expression coincides with a perpetuous activation of the selected protein). The type of mutation (KO or OE) has been previously related to the pathogenesis of intellectual disability and ASDs (Henle et al., 2011; Minshew and Williams, 2007; Sbacchi et al., 2010). The results of attractor analyses are presented in the Table 1. At first glance, the table shows that the number of attractors between Mutant and Default networks is different for all cases except the mutant Wave.

*Rho GTPases*. The Table 1 suggests opposite roles of RhoA and Rac/Cdc42 respectively on the activation of FE and Lam: OE of RhoA appears to exert a strong inhibitory effect on both nodes, whereas a KO of Rac or Cdc42 inactivates the nodes (Lam and FE respectively). The effects of in silico mutations of the three Rho GTPases were also assessed by time-series analysis. This analysis shows that the OE of RhoA strongly inhibits the activation of both Lam and FE nodes (see Fig. 5C and D). KO of the Rac node decrease to approximately 0% the activation of the Lam node (Fig. 5A). Finally, the KO of Cdc42 node had an inhibitory effect only on the FE node (Fig. 5B). These effects are evident in synchronous simulations.

**Table 1** Percentage of attractors with active (1) phenotypic nodes using knock-out (fixed 0) or overexpression (fixed 1) of the three Rho-GTPases and Wave nodes.

| | | *No. of attractors* | *FE** | *Lam** |
|---|---|---|---|---|
| Default network | | 72 | 21% | 42% |
| RAC | 0 | 32 | 0% | 0% |
| RhoA | 1 | 16 | 0% | 0% |
| Cdc42 | 0 | 36 | 0% | 42% |
| Wave | 0 | 72 | 21% | 0% |

*FE**, Filopodia elongation; *Lam**, Lamellipodia.

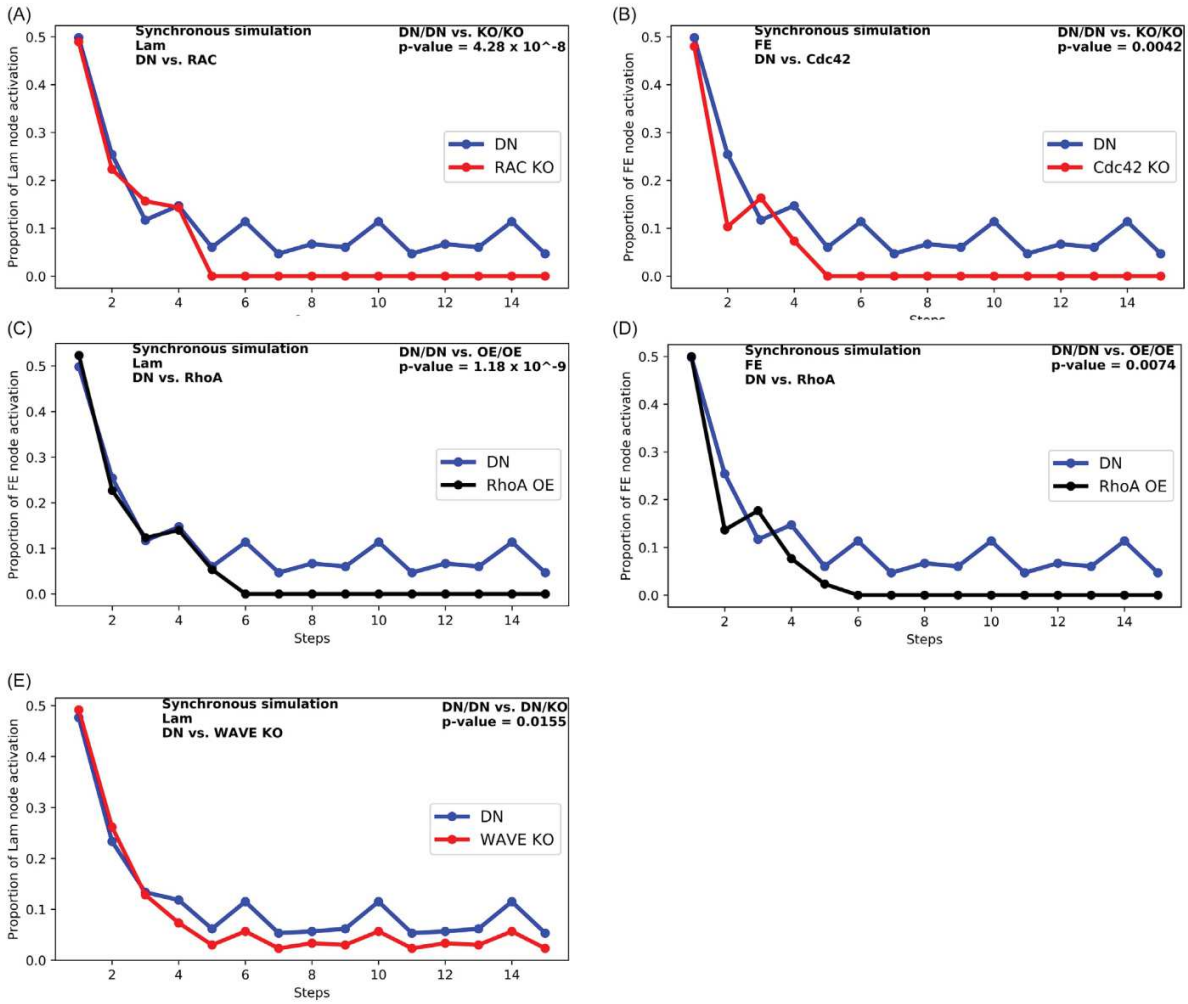



**Fig. 5** The profiles (graphs A–D) refer to simulations performed for the three Rho-GTPases homozygous genotypes. "DN/DN" = default (normal) network; "KO/KO" (knock-out) or "OE/OE" (over-expression, i.e., constitutive expression) = mutant networks. Graph (E) shows the profiles obtained by simulating a heterozygote genotype (DN/KO) for a deletion of CYFIP1 (Wave complex) gene detected in one autistic patient (Kawano et al., 2005; Leblond et al., 2012) compared to a DN/DN genotype. Global differences between each time-series pair of profiles (DN vs. mutant) are statistically significant as detected by the Brown's method (Poole et al., 2016) and application of the Bonferroni's corrections; threshold of significance = $p$-value $< 0.05$.

The essential roles played by the three Rho-GTPases in many functions related to cell movements strongly suggest that homozygous genotypes for amorphic mutations (corresponding to complete loss of function of the gene) would be lethal. Consequently, possible realistic biological scenarios simulated by permanently fixing the Rho GTPases node to 0 include: (i) homozygous for hypomorphic mutations, (ii) heterozygote for amorphic or hypomorphic mutations (respectively predicted to cause a complete loss of function or a reduced activity of the protein), (iii) effect of conditional KO/KO mutant mice, or (iv) epimutations (mutation affecting the expression but leave the DNA sequence unchanged) in specific neural cell types.

*CYFIP1 (Wave complex)*. The CYFIP1 gene encodes a protein that is part of the Wave complex (Kawano et al., 2005). According to the SFARI database this gene is classified in category 3 ("suggestive evidence") when associated to ASDs. There are however other reports associating the CYFIP1 gene with autism (Alvarez-Mora et al., 2016; Leblond et al., 2012; Toma et al., 2014; Waltes et al., 2014; Wang et al., 2015; van der Zwaag et al., 2010) and, in particular, Leblond et al. (2012) described a patient with a de novo deletion in chromosome 15 removing a few genes. To test the effect of this mutation in silico we generated a mutant (KO) BN by fixing the state of the Wave node to 0 and we then monitored the behavior of the two phenotypic nodes during simulation. The results of attractor analysis performed on the Wave-KO BN show a strong negative effect on the activation state of Lam node. Indeed, 100% of the attractors had an inactive (0) Lam node. These results were observed with both asynchronous and synchronous simulations (see Table 1 for the results of synchronous simulation). In this setting the in silico experiment just described is simulating a hypothetical patient with both CYFIP1 gene alleles deleted (i.e., homozygous). The effects of Wave/CYFIP1 KO were further assessed by time-series analysis. Since the patient studied by Leblond et al. (2012) bears a de novo Wave/CYFIP1 gene deletion it is reasonable to hypothesize that the patient was heterozygote for this mutation. To model the heterozygote state, we set the simulation with synchronous update of the nodes' states and observed the phenotypic nodes in their active state in the default (DN) and the CYFIP1 network (KO). The homozygote and heterozygote conditions were emulated by reporting the ratio of phenotypic nodes activity in DN compared to DN network and in KO compared to KO network respectively. This approach revealed a reduction of approximately 50% in the rate of activation of the Lam node between the normal (DN/DN) and heterozygote (DN/KO) simulated genotypes (see Fig. 5E).

In summary, the simulations performed with mutant networks suggest that, in pathological conditions, the cellular phenotype affected by mutations or epimutations affecting four tested genes is the motility of the GC and more specifically, the formation of filopodia and/or lamellipodia. It is important to highlight that these results are consistent with previous studies (reviewed in Henle et al., 2011) indicating that in intellectual disability and ASDs the RhoA pathway is hyperactive, while the Rac and Cdc42 pathways are consistently hypoactive.

## Discussion and Concluding Remarks

In this study we investigated the molecular dynamics underlying the behavior of the neuronal GC during AGG by approaching the problem from a systemic perspective. For this reason and to cope with the increasing amount of available data, we applied the strategies integrated in Systems Biology approach. As a first step, we have used dynamic Boolean modeling to investigate the functional relationship of a subset of 27 proteins to GC behavior. In human species, more than 350 proteins have been related to the regulation of this highly complex process. By inspecting the connections of these proteins within and among the signaling pathways they are involved into, it was observed that several guidance signals, i.e., netrins, ephrins, slits, and semaphorins, converge to the three Rho GTPases (i.e., Rac, Cdc42, RhoA) through GC membrane receptors and signal transduction pathways, hence constituting a necessary relay station of many upstream signaling pathways. Indeed, these three GTPases induce changes in cytoskeletal organization that determine the direction in which the GC will turn to (Human axon guidance from Kyoto Encyclopedia of Genes and Genomes (KEGG hsa04360)). Based on these observations, a reduced version of the GC network was constructed by removing signaling nodes converging to the three Rho-GTPases. Furthermore, intermediate proteins (for example, Rock from the sequence RhoA > Rock > Limk) have been also removed due to the assumption that they do not regulate nor are regulated by any other element of the network. With this pruning procedure, it becomes evident that the network focuses on nodes that are critical for the propagation of signals coming from GC plasma membrane. Compared to other available networks accounting for AG, our network significantly differ due to the reduction procedure and the contextualization of nodes and edges to the neuronal development. The manual curation ensures that 100% of the nodes are neuron-specific and more than 80% of the nodes have been have been proven in rodent cortical neurons (based of our bibliomic analysis).

Despite of the limited number of proteins and phenotypes included, the GC network combined with the corresponding Boolean model constitute a powerful tool for investigating and unveiling new details about AGG and GC behavior. The initial results here presented are promising, and therefore it would be of interest to improve the current version of the network by including additional elements. The addition of phenotypic nodes corresponding to cellular behaviors (e.g., axon branching, neurite extension, etc.) necessitate to be properly connected to the current network hence including their related proteins. For example, the introduction of attractive and/or repulsive extracellular cues (e.g., netrin, slit, ephrin, semaphorin) as upstream signaling cascade of the three Rho-GTPases can potentially unveil new relation between signals exchange driving the cellular behavior. Despite of the lack of the upstream signaling in the current version of the network, we aimed to reproduce these cues (data not shown) by using the "fixedGenes()" function to model the combined effects of repulsive and attractive extracellular signals on filopodia and lamellipodia formation. To simulate the combination of molecular cues from upstream signaling cascade, eight possible binary configurations of the three Rho-GTPases nodes (23) represent the triggering of the signaling cascade. These configurations recapitulate all possible signals arriving from the extracellular cues interacting with their specific receptors. For example, a stable and/or long-lasting attractive signal mediated by the interaction of Netrin with DCC would correspond to the following configuration: Rac = 1 AND Cdc42 = 1 AND RhoA = 0. Similarly, the contemporary presence in the Extracellular Matrix (ECM) of two cues of opposite signs (e.g., Sema3A/Plexin plus Netrin/DCC) would correspond in our simulation to only one configuration that is Rac = 1 AND Cdc42 = 1 AND RhoA = 1. In this simulation, the state of the remaining nodes is changing according to; (i) the (randomly chosen) initial values of nodes, (ii) the updating scheme (synchronous/asynchronous), (iii) the transition rules used.

Since Boolean models rely on binary states, quantitative aspects cannot be included within the current version of the model. However, by transforming the model either into a multi-level logic model or to a fuzzy logic Boolean model, we should circumvent the afore mentioned limitations. If successful, these improvements shall increase the predictive power of our model.

The different simulation paradigms used above can also have significant medical relevance in terms of relating biochemical networks with pathological conditions as neurodevelopmental disorders or autism. Indeed, as we have shown, simulations performed with mutant networks represent an important potential application of the model that allows for an understanding at a system level of the effect of mutations. In fact, in neurodevelopmental disorders, conventional approaches for assessing the functional effects of a gene mutation usually rely on the analysis of how the mutation affects the corresponding protein (i.e., changes in structure and function, degradation, or changes in its level of expression). Following this approach, the dynamic effects of the resulting changes on the interactions with other proteins are mostly ignored or evaluated through the analysis of static pathways. Our model offers an additional perspective and fills this gap thanks to the systemic approach. Based on these encouraging results, we propose the use of this version of the Boolean model of AGG as a valuable tool to support current methods used for the prediction of functional and/or phenotypic effects of single as well as simultaneous mutation of multiple genes. This perspective is especially relevant in a clinical context. More precisely, we claim that in silico evaluation of the combined effects of multiple mutations identified by exome sequencing through such a powerful tool might constitute a significant twist in the approach to a single patient affected by multifactorial neurodevelopmental disorders caused by a defective AGG such as ASDs (Box 2).

**Box 2 Future perspectives for cellular and molecular Boolean modeling of ASDs**

Boolean networks and their computational implementations is an effective and flexible technology that can be easily adapted to tackle the challenges of modeling a complex dynamic system like ASDs. This is indeed a developmental disorder in constant change, from cellular to environmental changes. In addition, these changes can occur over precisely-defined neurodevelopmental time points or stages and at different timescales. In an attempt to cope with this complexity, we plan to develop new expanded and improved versions of the Boolean algorithms. Hereafter, we briefly outline in which directions we will invest our future efforts. *Boolean Networks (BNs) for other neurodevelopmental processes: networks and supernetworks.* The size and topology of BNs can be easily modified by adding or removing nodes (representing proteins, genes, coding and non-coding RNAs, other molecular species) and/or interactions (arcs) to simulate the dynamics underlying other cellular processes beside axon guidance. These processes may occur at the same or different neurodevelopmental time points or stages and include, among others: neuronal proliferation, migration, apoptosis, synaptogenesis, myelination. The latter processes can be modeled either as distinct molecular networks or as a single integrated supernetwork (e.g., Schlatter et al., 2012) reconstructed thanks to shared nodes usable as bridging points between different networks (e.g., UNC-6/netrin and LIN-44/Wnt participate in axon guidance and synaptogenesis, Shen and Cowan, 2010). *Microenvironments and initial conditions.* The particular microenvironments (intracellular/extracellular) known or assumed to influence one or more processes (e.g., presence/absence of extracellular cues influencing axon guidance, neuronal migration or synaptogenesis) can be modeled by choosing the specific initial conditions. A particular initial condition is the ON/OFF configuration owned by each node in the network before the start of simulation. *Timescales and timedelays.* The association of attractors with specific phenotypes must take into account possible differences existing in timescales between molecular and cellular events. Similarly, different timescales may also occur between distinct molecular events—on one side—(e.g., the binding of Netrins to DCC must precede the activation of Rac) or between distinct cellular events—on the other side (e.g., axon guidance precedes synaptogenesis).

Different timescales can be taken into account during simulation by introducing timedelays to the logical functions. For instance, a specific function is provided in *BoolNet* by which processes can be assigned to different timescales. These timescales are constants that specify in which state a certain node can become active. Specifically, simulating a network at timescale $t = x$ means that all interactions with a timescale constant $t = x$ are considered, but interactions with timescale constant $t \neq x$ are omitted. In other words, it is like a different subnetwork is created at each time step. Such an approach has been successfully used for the modeling of apoptosis (Schlatter et al., 2009).

*Potential contributions of BN modeling for understanding the actual development of ASDs.* The approaches outlined above can also be used to model pathogenic processes. This has important implications to understand the actual development of ASDs. In past years, structural and functional MRI studies have shown that a prominent and consistent feature of these disorders is an altered pattern of neuronal connectivity: both low connectivity between distant cortical areas and local hyperconnectivity have been described (Holiga et al., 2019). Recently, as mentioned in the Introduction, omic and clinical studies have suggested potentially critical genetic, epigenetic or environmental factors that acting individually or in combination are likely to generate an altered pattern of connectivity. Interestingly, functional annotation studies have shown that among the best molecular candidates stand out components of signaling pathways regulating synaptogenesis, axon guidance and other neurodevelopmental processes all of which we aim to model by BNs (see above). Unfortunately, in most cases the actual causal involvement of a gene mutation in ASDs is based on circumstantial evidence because direct functional in vitro or in vivo experimental data are lacking. It is precisely for this reason that Boolean models can be very useful. For example, suppose that exome sequencing has revealed in a patient the presence in its DNA of one or more mutations in different genes. The "fixGenes" function of *BoolNet* can then be used to test the effect of single mutations or even groups of mutations on the dynamics of Boolean networks. These networks are reconstructed from signaling pathways regulating one or more neurodevelopmental processes. Finally, by monitoring the state of phenotypic nodes (ON/OFF) at the end of a simulation, i.e., when an attractor is reached, it should be possible at least in silico to identify which mutant proteins and initial conditions are most likely involved in the disruption of one or more neurodevelopmental process. In other words, this diagnostic procedure should lead to the identification of "pathogenic attractors," i.e., an easy term we are using here to define this type of in silico dynamic biomarker, analogous to "signature," a term that is often used to define the transcriptomic profile associated to a specific form of, e.g., cancer cells or a particular differentiated cell type. Other interesting applications of in silico mutagenesis with BNs concern the possibility of discovering targets for therapeutic drugs. Briefly, such potential targets would coincide with the nodes that when knocked-out or activated permanently during the simulation of a mutant network or supernetwork are able to revert pathogenic attractors to normal attractors.

## References


Abrahams BS, Arking DE, Campbell DB, et al. (2013) SFARI gene 2.0: A community-driven knowledgebase for the autism spectrum disorders (ASDs). *Molecular Autism* 4: 36.

Albert R and Othmer HG (2003) The topology of the regulatory interactions predicts the expression pattern of the segment polarity genes in Drosophila melanogaster. *Journal of Theoretical Biology* 223: 1–18.

Albert R and Robeva R (2015) Signaling networks. In: *Algebraic and discrete mathematical methods for modern biology*, pp. 65–91. Elsevier.

Aletti G and Causin P (2008) Mathematical characterisation of the transduction chain in growth cone pathfinding. *IET Systems Biology* 2: 150.

Alvarez-Mora MI, Calvo Escalona R, Puig Navarro O, et al. (2016) Comprehensive molecular testing in patients with high functioning autism spectrum disorder. *Mutation Research* 784–785: 46–52.

Banerjee-Basu S and Packer A (2010) SFARI gene: An evolving database for the autism research community. *Disease Models & Mechanisms* 3: 133–135.

Charron F (2005) Novel brain wiring functions for classical morphogens: A role as graded positional cues in axon guidance. *Development* 132: 2251–2262.

Codina-Solà M, Rodríguez-Santiago B, Homs A, et al. (2015) Integrated analysis of whole-exome sequencing and transcriptome profiling in males with autism spectrum disorders. *Molecular Autism* 6: 21.

Cohen IR and Harel D (2007) Explaining a complex living system: Dynamics, multi-scaling and emergence. *Journal of the Royal Society Interface* 4: 175–182.

Dent EW, Gupton SL, and Gertler FB (2011) The growth cone cytoskeleton in axon outgrowth and guidance. *Cold Spring Harbor Perspectives in Biology* 3: a001800.

Engle EC (2010) Human genetic disorders of axon guidance. *Cold Spring Harbor Perspectives in Biology* 2: a001784.

Fischer HP (2008) Mathematical modeling of complex biological systems: From parts lists to understanding systems behavior. *Alcohol Research & Health* 31: 49–59.

Forbes EM, Thompson AW, Yuan J, and Goodhill GJ (2012) Calcium and cAMP levels interact to determine attraction versus repulsion in axon guidance. *Neuron* 74: 490–503.

Fumiã HF and Martins ML (2013) Boolean network model for cancer pathways: Predicting carcinogenesis and targeted therapy outcomes. *PLoS One* 8: e69008.

Gasperini RJ, Pavez M, Thompson AC, et al. (2017) How does calcium interact with the cytoskeleton to regulate growth cone motility during axon pathfinding? *Molecular and Cellular Neurosciences* 84: 29–35.

Gilman SR, Iossifov I, Levy D, et al. (2011) Rare de novo variants associated with autism implicate a large functional network of genes involved in formation and function of synapses. *Neuron* 70: 898–907.

Griswold AJ, Dueker ND, Van Booven D, et al. (2015) Targeted massively parallel sequencing of autism spectrum disorder-associated genes in a case control cohort reveals rare loss-of-function risk variants. *Molecular Autism* 6: 43.

Hall A and Lalli G (2010) Rho and Ras GTPases in axon growth, guidance, and branching. *Cold Spring Harbor Perspectives in Biology* 2: a001818.

Henle SJ, Wang G, Liang E, et al. (2011) Asymmetric PI(3,4,5)P3 and Akt signaling mediates chemotaxis of axonal growth cones. *The Journal of Neuroscience* 31: 7016–7027.

Hetmanski JHR, Schwartz JM, and Caswell PT (2018) Rationalizing Rac1 and RhoA GTPase signaling: A mathematical approach. *Small GTPases* 9: 224–229.

Holiga Š, Hipp JF, Chatham CH, Garces P, Spooren W, D'Ardhuy XL, Bertolino A, Bouquet C, Buitelaar JK, Bours C, et al. (2019) Patients with autism spectrum disorders display reproducible functional connectivity alterations. *Science Translational Medicine* 11: eaat9223.

Huang S (1999) Gene expression profiling, genetic networks, and cellular states: An integrating concept for tumorigenesis and drug discovery. *Journal of Molecular Medicine* 77: 469–480.

Hussman JP, Chung RH, Griswold AJ, et al. (2011) A noise-reduction GWAS analysis implicates altered regulation of neurite outgrowth and guidance in autism. *Molecular Autism* 2: 1.

Ingalls BP (2013) *Mathematical modeling in systems biology: An introduction.* Cambridge, MA: MIT Press.

Irurzun-Arana I, Pastor JM, Trocóniz IF, and Gómez-Mantilla JD (2017) Advanced boolean modeling of biological networks applied to systems pharmacology. *Bioinformatics* 33(7): 1040–1048.

Iyer AN, Bellon A, and Baudet ML (2014) microRNAs in axon guidance. *Frontiers in Cellular Neuroscience* 8: 78.

Kalkman HO (2012) Potential opposite roles of the extracellular signal-regulated kinase (ERK) pathway in autism spectrum and bipolar disorders. *Neuroscience & Biobehavioral Reviews* 36: 2206–2213.

Kauffman SA (1993) *The origins of order: Self-organization and selection in evolution.* New York: Oxford University Press.

Kawano Y, Yoshimura T, Tsuboi D, et al. (2005) CRMP-2 is involved in kinesin-1-dependent transport of the Sra-1/WAVE1 complex and axon formation. *Molecular and Cellular Biology* 25: 9920–9935.

Kitano H, Funahashi A, Matsuoka Y, and Oda K (2005) Using process diagrams for the graphical representation of biological networks. *Nature Biotechnology* 23: 961–966.

Kolodkin AL and Tessier-Lavigne M (2011) Mechanisms and molecules of neuronal wiring: A primer. *Cold Spring Harbor Perspectives in Biology* 3: a001727.

Koyama R and Ikegaya Y (2018) The molecular and cellular mechanisms of axon guidance in mossy fiber sprouting. *Frontiers in Neurology* 9: 382.

Krottje JK and van Ooyen A (2007) A mathematical framework for modeling axon guidance. *Bulletin of Mathematical Biology* 69: 3–31.

Leblond CS, Heinrich J, Delorme R, et al. (2012) Genetic and functional analyses of SHANK2 mutations suggest a multiple hit model of autism spectrum disorders. *PLoS Genetics* 8: e1002521.

Li F, Long T, Lu Y, Ouyang Q, and Tang C (2004) The yeast cell-cycle network is robustly designed. *Proceedings of the National Academy of Sciences of the United States of America* 101: 4781–4786.

Lowery LA and Van Vactor DV (2009) The trip of the tip: Understanding the growth cone machinery. *Nature Reviews. Molecular Cell Biology* 10: 332–343.

McFadden K and Minshew NJ (2013) Evidence for dysregulation of axonal growth and guidance in the etiology of ASD. *Frontiers in Human Neuroscience* 7: 671.

Minshew NJ and Williams DL (2007) The new neurobiology of autism: Cortex, connectivity, and neuronal organization. *Archives of Neurology* 64: 945–950.

Padmanabhan P and Goodhill GJ (2018) Axon growth regulation by a bistable molecular switch. *Proceedings of the Royal Society B: Biological Sciences* 285: 2017–2618.

Piper M, van Horck F, and Holt C (2007) The role of cyclic nucleotides in axon guidance. In: Bagnard D (ed.) *Axon growth and guidance*, pp. 134–143. New York, NY: Springer.

Poole W, Gibbs DL, Shmulevich I, Bernard B, and Knijnenburg TA (2016) Combining dependent P-values with an empirical adaptation of Brown's method. *Bioinformatics* 32: i430–i436.

Roccasalvo IM, Micera S, and Sergi PN (2015) A hybrid computational model to predict chemotactic guidance of growth cones. *Scientific Reports* 5: 11430.

Sbacchi S, Acquadro F, Calò I, Calì F, and Romano V (2010) Functional annotation of genes overlapping copy number variants in autistic patients: Focus on axon pathfinding. *Current Genomics* 11: 136–145.

Schlatter R, Schmich K, Avalos Vizcarra I, Scheurich P, Sauter T, Borner C, Ederer M, Merfort I, and Sawodny O (2009) ON/OFF and beyond—A Boolean model of apoptosis. *PLoS Computational Biology* 5: e1000595.

Schlatter R, Philippi N, Wangorsch G, Pick R, Sawodny O, Borner C, Timmer J, Ederer M, and Dandekar T (2012) Integration of Boolean models exemplified on hepatocyte signal transduction. *Briefings in Bioinformatics* 13: 365–376.

Shen K and Cowan CW (2010) Guidance molecules in synapse formation and plasticity. *Cold Spring Harbor Perspectives in Biology* 2: a001842.

Stoeckli ET (2018) Understanding axon guidance: Are we nearly there yet? *Development* 145: dev151415.

Suda S, Iwata K, Shimmura C, et al. (2011) Decreased expression of axon-guidance receptors in the anterior cingulate cortex in autism. *Molecular Autism* 2: 14.

Tangsuwansri C, Saeliw T, Thongkorn S, et al. (2018) Investigation of epigenetic regulatory networks associated with autism spectrum disorder (ASD) by integrated global LINE-1 methylation and gene expression profiling analyses. *PLoS One* 13: e0201071.

Tessier-Lavigne M (2002) Wiring the brain: The logic and molecular mechanisms of axon guidance and regeneration. *Harvey Lectures* 98: 103–143.

Tessier-Lavigne M and Goodman CS (1996) The molecular biology of axon guidance. *Science* 274: 1123–1133.

Toma C, Torrico B, Hervás A, et al. (2014) Exome sequencing in multiplex autism families suggests a major role for heterozygous truncating mutations. *Molecular Psychiatry* 19: 784–790.
Vallabhajosyula RR and Raval A (2010) Computational modeling in systems biology. In: Yan Q (ed.) *Systems biology in drug discovery and development*, pp. 97–120. Totowa, NJ: Humana Press.
van der Zwaag B, Staal WG, Hochstenbach R, et al. (2010) A co-segregating microduplication of chromosome 15q11.2 pinpoints two risk genes for autism spectrum disorder. *American Journal of Medical Genetics* 153B: 960–966.
Verleysen M (2013) Université catholique de Louvain, Katholieke Universiteit Leuven. In: *Proceedings/21st European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning, ESANN 2013: Bruges, Belgium, April 24-25-26, 2013. Ciaco, Louvain-la-Neuve.*
Waltes R, Duketis E, Knapp M, et al. (2014) Common variants in genes of the postsynaptic FMRP signalling pathway are risk factors for autism spectrum disorders. *Human Genetics* 133: 781–792.
Wang J, Tao Y, Song F, et al. (2015) Common regulatory variants of CYFIP1 contribute to susceptibility for autism spectrum disorder (ASD) and classical autism. *Annals of Human Genetics* 79: 329–340.
Wang Z, Li P, Wu T, et al. (2018) Axon guidance pathway genes are associated with schizophrenia risk. *Experimental and Therapeutic Medicine* 16(6): 4519–4526.
Zamboni V, Jones R, Umbach A, et al. (2018) Rho GTPases in intellectual disability: From genetics to therapeutic opportunities. *International Journal of Molecular Sciences* 19: 1821.

## Relevant Websites

http://geneontology.org/—The Gene Ontology Resource.
https://www.sfari.org/2013/03/12/rho-family-of-enzymes-at-crossroads-of-autism/—SFARI.
https://gene.sfari.org/database/humangene/—SFARI Gene.
http://wannovar.wglab.org/—wANNOVAR.
http://genetics.bwh.harvard.edu/pph2/—PolyPhen-2.
https://www.r-project.org/—The R Project.
https://cran.r-project.org/web/packages/BoolNet/index.html—BoolNet.
https://www.genome.jp/dbget-bin/www_bget?pathway+hsa04360—KEGG Axon Guidance pathway.