SIGNALSENSE ATELIER · EVOLUTIONARY INTELLIGENCE MONOGRAPH

Running to Remain

Leigh Van Valen, the Red Queen Hypothesis, and the Coevolutionary Future

Prepared for Joaquim A. Machado · 12 March 2026

Developed with OpenAI GPT-5.4 Thinking

A deep-dive essay on evolutionary theory, coevolution, artificial intelligence, and adaptive agents

Thesis

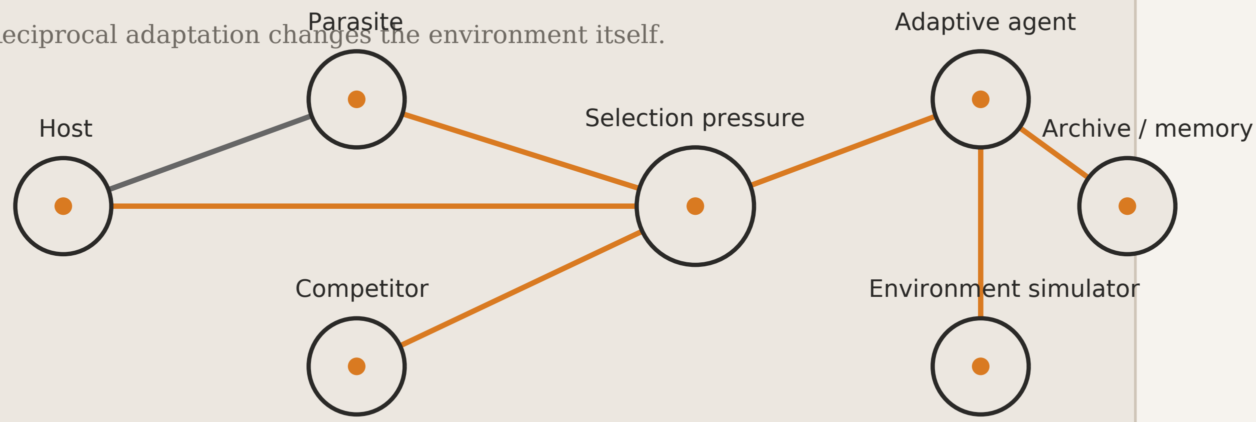
The Red Queen remains one of the most useful concepts for thinking about both living evolution and machine adaptation, because it treats the environment as a moving ecology of opponents, allies, and constraints—not as a static backdrop.

Executive Abstract

Leigh Van Valen's Red Queen hypothesis remains one of the most fertile ideas in evolutionary theory because it shifted attention from organisms adapting to a fixed world toward lineages adapting inside worlds that other lineages are also changing. That move now resonates strongly with contemporary host–parasite genomics, macroevolutionary theory, competitive co-evolution, self-play, LLM-agent training, and adaptive safety. This monograph revisits Van Valen's original insight, examines how current evolutionary biology has refined it, and argues that the Red Queen is increasingly becoming not only a biological explanation but also a design principle for artificial intelligence, artificial life, and resilient adaptive agents.

Figure 1. From Red Queen biology to adaptive-agent design

Reciprocal adaptation changes the environment itself.



The diagram condenses the core translation developed in this report: host–parasite and competitor dynamics in biology reappear in AI as adaptive-agent ecologies, where archives, memory, and evolving evaluators are required to distinguish genuine progress from cyclical forgetting.

Why this matters now

Van Valen's decisive move was to treat other evolving systems as part of the environment. That move now links fossil survivorship, pathogen-driven diversity, multi-agent reinforcement learning, self-evolving LLM agents, and adaptive safety.

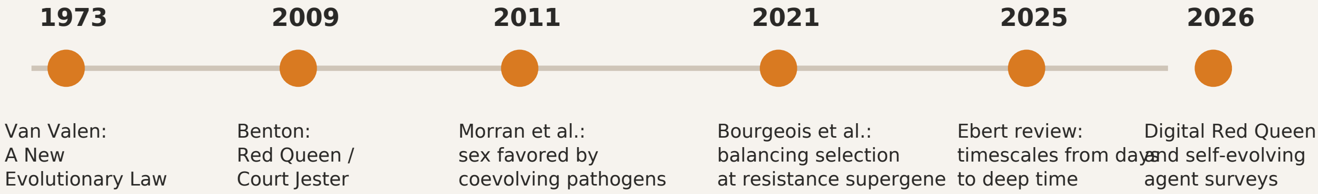
1. Van Valen’s original break

Leigh Van Valen’s 1973 paper, A New Evolutionary Law, proposed a startling claim: survivorship curves from the fossil record suggested that the probability of extinction within many clades remains approximately constant through time, even after lineages have already persisted for long periods [1]. The point was not that evolution fails to improve organisms, but that improvement is relative. A lineage survives not because it has reached some permanent optimum, but because it keeps pace with antagonists, competitors, prey, partners, and pathogens that are also changing.

That insight became the Red Queen hypothesis. The literary source was Lewis Carroll’s Red Queen, who tells Alice that one must run as fast as possible simply to stay in the same place. Van Valen converted that image into an evolutionary principle: fitness is not best understood as an absolute score, but as a moving relationship inside an ecological web. Solé’s retrospective emphasizes the same point—species do not merely evolve against a background; they also coevolve with other species that constitute a major part of the environment itself [2].

> *“Species do not merely evolve: they also coevolve with other species.”*

This was conceptually decisive because it linked paleobiology to ecology, ecology to genetics, and genetics to the logic of arms races. It also anticipated later work on complex adaptive systems, where local improvement can coexist with persistent instability because the target of adaptation is itself adapting [1][2].



The timeline shows why the Red Queen survives conceptually. It begins as a paleobiological inference, gains mechanistic force in ecology and genomics, and then migrates into artificial-life and agentic-AI research.

2. What current evolutionary biology says

Current evolutionary biology supports the Red Queen most strongly in host–parasite systems. These are the settings where reciprocal selection is fast, negative frequency-dependent selection can be measured, and genomic traces of the struggle are often visible. A central prediction is that common host genotypes become easier for parasites to exploit, making rare host variants advantageous until they too become common. This generates continual turnover instead of permanent victory.

Experimental evolution has made that logic unusually concrete. Morran and colleagues showed that reciprocal coevolution between Caenorhabditis elegans and Serratia marcescens increased outcrossing, while obligately selfing lineages went extinct under pathogen coevolution [3]. Papkou and collaborators then dissected the genomic basis of rapid host–pathogen coevolution, finding distinct selective processes in host and pathogen during controlled experimental Red Queen dynamics [4].

Genomic work in the Daphnia magna–Pasteuria ramosa system has deepened the picture. Bourgeois and colleagues found signatures of long-term balancing selection at a resistance supergene, including high diversity, older coalescence times, and reduced differentiation relative to the genomic background [5]. Dieter Ebert’s 2025 review extends this logic across timescales—from direct observations spanning days or years to archived material, postglacial patterns, and deep phylogeographic time—arguing that the evidence strongly supports the Red Queen as an explanation for extraordinary disease-gene diversity [6].

At macroevolutionary scales, however, the modern picture is more plural than early Red Queen discussions sometimes implied. Benton’s Red Queen/Court Jester synthesis and Condamine’s work on Apollo butterflies both show that biotic interactions matter deeply, but so do climate, geography, and mass extinction. The contemporary lesson is not that the Red Queen was wrong, but that it operates as one major causal regime inside a broader multiscale theory of evolution [7][8].

Translational map

Negative frequency-dependent selection: rare host genotypes gain temporary advantage because parasites adapt to common forms [5][6]. In AI, diverse policy populations and cross-play reduce exploitation of a single dominant policy.

Persistence of sex / recombination: reciprocal pathogen pressure can favor outcrossing over selfing [3]. In AI, population-based training and recombination-like model mixing can increase robustness.

Coevolutionary time-lags: adaptation unfolds through delayed reciprocal response rather than instant equilibrium [4][6]. Historical evaluation matters: an agent that wins now may fail against earlier or unseen opponents [9].

Multiscale causation: Red Queen and Court Jester factors interact across ecological and geological scales [7][8]. Likewise, agent performance depends both on adversaries and on exogenous shifts such as deployment context, user behavior, or platform change.

3. From host–parasite arms races to adaptive-agent ecologies

Van Valen matters for artificial intelligence because multi-agent learning turns the environment into a moving target. In self-play and competitive co-evolution, agents are not optimized against a fixed world; they are optimized against other learners whose improvements alter the task distribution. That is a digital Red Queen.

Nolfi and Pagliuca’s 2025 systematic comparison makes the engineering problem explicit: competitive co-evolution exposes agents to continuously varying environmental conditions, but it can also produce endless limit cycles in which strategies are discovered, forgotten, and rediscovered. Their remedy is historically informed evaluation—archives of opponents, diversity-aware selection, and metrics for distinguishing local progress from genuine global progress [9].

Recent LLM-agent research has started to operationalize the same intuition. Multi-Agent Evolve organizes a proposer, solver, and judge in a co-evolving loop; the authors report average benchmark improvement from this reciprocal training architecture [10]. GenEnv makes the environment itself adaptive: an Environment LLM generates tasks at the edge of the agent’s capability, forming a difficulty-aligned co-evolutionary curriculum rather than a static dataset [11]. The 2026 Digital Red Queen work takes the biological analogy even more literally, using self-play in Core War to study open-ended adversarial program evolution under changing objectives [12].

The same logic is entering AI safety. The R^2AI position paper argues that open-ended worlds will continue producing novel threats faster than fixed safety checklists can absorb them, and therefore proposes “safe-by-coevolution”: continual adversarial, adaptive, internally coupled safety processes rather than episodic external patching [13]. The 2026 survey of self-evolving agents likewise treats co-evolution, memory, and evaluation as core ingredients of the next generation of adaptive systems [14].

Engineering restatement

The Red Queen translation into AI is strongest when we stop asking whether an agent solves a fixed task and start asking whether it remains competent while the task generator, adversary, evaluator, and safety pressures co-adapt.

4. Design principles for adaptive agents

1. Diversity is not redundancy.

Red Queen biology preserves variation because rare genotypes often gain the advantage. In AI, archives, ensembles, cross-play populations, and heterogeneous memories are not ornamental extras; they are anti-exploitation infrastructure.

2. Evaluation must be ecological, not merely pointwise.

A strong agent is not one that beats the current opponent of the day; it is one that remains competent across shifting ecologies of opponents, tasks, and contexts. Static leaderboards hide local overfitting.

3. Memory is an anti-forgetting organ.

Competitive co-evolution easily produces cycles of discovery and loss. Memory buffers, archived opponents, retained lineages, and rollback mechanisms function like immunological or phylogenetic memory against regression.

4. Curriculum should be endogenous.

The most efficient challenges are often generated by the evolving system itself. Co-evolving adversaries, judges, or environment models naturally place an agent near its adaptive frontier.

5. Safety must evolve with capability.

If capability is open-ended but safety is static, safety becomes obsolete by design. Red Queen thinking implies closed-loop challenge, critique, rollback, and recovery mechanisms.

5. Conclusion: from evolution theory to machine ecology

The deeper philosophical lesson is that intelligence should not be modeled only as optimization over a fixed objective. In the Red Queen frame, intelligence is persistence under reciprocal adaptation. This makes the concept especially relevant for artificial life, agentic AI, embodied systems, robotics, cybersecurity, and any setting in which the surrounding world contains other learners, other strategies, or other evolving constraints. The unit of analysis shifts from the isolated agent to the coevolutionary loop.

That shift may explain why many contemporary AI systems look impressive on closed tasks yet brittle in open settings. They have been trained to solve problems, not to survive ecologies of changing problems. Van Valen’s deeper contribution was to show that survival in such ecologies depends less on absolute excellence than on continual revision under relation. For AI, the practical corollary is stark: the next frontier is not simply stronger models, but better coevolutionary regimes.

Final assessment

Van Valen's real legacy is methodological. He taught us to look for adaptation not in isolation, but in feedback-rich ecologies where survival depends on continual revision. That insight is as relevant to pathogens and genomes as it is to agentic AI.

Prepared in SignalSense Atelier style for Joaquim A. Machado, with OpenAI GPT-5.4 Thinking acknowledged as the AI analytical and editorial collaborator for this edition.

References

[1] Van Valen, L. (1973). A New Evolutionary Law. Evolutionary Theory, 1, 1–30.

[2] Solé, R. (2022). Revisiting Leigh Van Valen's "A New Evolutionary Law" (1973). Biological Theory, 17, 120–125.

[3] Morran, L. T., Schmidt, O. G., Gelarden, I. A., Parrish, R. C., & Lively, C. M. (2011). Running with the Red Queen: host–parasite coevolution selects for biparental sex. Science, 333(6039), 216–218.

[4] Papkou, A., Guzella, T., Yang, W., et al. (2019). The genomic basis of Red Queen dynamics during rapid reciprocal host–pathogen coevolution. PNAS, 116(3), 923–928.

[5] Bourgeois, Y., Fields, P. D., Bento, G., & Ebert, D. (2021). Balancing selection for pathogen resistance reveals an intercontinental signature of Red Queen coevolution. Molecular Biology and Evolution, 38(11), 4918–4933.

[6] Ebert, D. (2025). The Red Queen and the timescale of antagonistic coevolution: parasite selection for genetic diversity. Annual Review of Genetics, 59, 215–236.

[7] Benton, M. J. (2009). The Red Queen and the Court Jester: species diversity and the role of biotic and abiotic factors through time. Science, 323(5915), 728–732.

[8] Condamine, F. L., Rolland, J., Höhna, S., Sperling, F. A. H., & Sanmartín, I. (2018). Testing the role of the Red Queen and Court Jester as drivers of the macroevolution of Apollo butterflies. Systematic Biology, 67(6), 940–964.

[9] Nolfi, S., & Pagliuca, P. (2025). Global progress in competitive co-evolution: a systematic comparison of alternative methods. Frontiers in Robotics and AI, 11:1470886.

[10] Chen, Y., et al. (2025/2026). Multi-Agent Evolve: LLM Self-Improve through Multi-Agent Co-evolution. OpenReview / preprint version accessed March 2026.

[11] Guo, J., et al. (2025). GenEnv: Difficulty-Aligned Co-Evolution Between LLM Agents and Environment Simulators. arXiv preprint 2512.19682.

[12] Kumar, A., Bahlous-Boldi, R., Sharma, P., Isola, P., Risi, S., Tang, Y., & Ha, D. (2026). Digital Red Queen: Adversarial Program Evolution in Core War with LLMs. arXiv preprint 2601.03335.

[13] Sun, Y., Wang, X., Fu, J., Lu, C., & Zhou, B. (2025). R²AI: Towards Resistant and Resilient AI in an Evolving World. Position paper / preprint.

[14] Gao, H., et al. (2026). A Survey of Self-Evolving Agents. TMLR survey version accessed March 2026.

Joaquim Machado & OpenAI GPT 5.4 Agent/Thinking