THE CRUCIBLE

2026

Table of Contents

Introduction to the Study....4
- Time and Space....5
- Let There be Light....5
- The Players....6
- Life and Death....7
- The Game....7
- Complexity....8
- The World Engine....8
- The Grand Experiment....9
- What Is Being Studied....11
- How This Study Should Be Read....11

Project 1: Baseline Crucible....13
- Part I - Project Plan....14
- Terms Used in This Report....14
- Part II - Set A: Baseline World....16
- Part III - Set B: Bottleneck World....18
- Part IV - Set C: Duel World....20
- Part V - Climate Severity Comparison....22
- Part VI - Project 1 Synthesis....24

Project 2: Competition....26
- Part I - Research Plan and Execution Design....27
- Central Research Frame....27
- Active Tribe Roster....27
- Primary Research Questions....29
- Working Hypotheses....29
- Part II - Data Summary and Outcome Analysis....32
- Executive Summary....32
- Final Conclusion....43

Project 3: Evolution....44
- Part I - Research Plan....45
- Primary Research Questions....47
- Working Hypotheses....48
- Part II - Outcome Analysis....52
- Executive Finding....52
- Final Interpretive Thesis....59

Project 4: Bombs....60
- Introduction and Study Design....61
- Terms Used in This Report....62
- Research Questions and Measures....63
- Phase 1 - Aerosol Demasking....68
- Phase 2 - Arctic Methane Release....70

Phase 3 - Nuclear Meltdown....73
Phase 4 - All Bombs....81
Cross-Phase Synthesis....88
Epilogue....89
Epilogue Research Overview....90
Research Plan and Working Objectives....90
Obsidian Reference Analysis....93
Ares Design Theory....96
Ares Field Test....100
Natural Network Collapse....103
Extended 80,000-Pulse Horizon....106
Reservoir Calibration....110
Final Climate Reservoir Projection....112
Closing Synthesis....117
Closing Frame and Five-Project Arc....118
General Results and Adaptation....119
Limits, Methodological Meaning, and Final Reflection....121
Appendices....124
Introduction, Project 1, Closing, and Epilogue Appendices....125
Addendum: Lessons for a Future Crucible Study....144

THE CRUCIBLE

Introduction to the Study

Opening frame, world engine, and study design

The explosion of complex life over millennia of simulated time starts with humble beginnings. Scattered tribes of wandering nomads are dropped into a simulated frozen wasteland 1000 pulses before the dawn of the Holocene.

Eleven large language models were each asked to design a tribe of 20 agents and to co-design a second tribe with the guidance of a biological contributor. They were given a set of 33 behavioural parameters to be configured for each agent as they saw fit. The LLM's were told that their founding agents would be the first generation of participants in an epoch-spanning simulation of life and death in a vibrant world of competition, cooperation, geography, climate and thermodynamic chaos. They were told that the goals and strategies of their design philosophy were entirely up to them. This was important because the crucial first step of the experiment was to give its players agency. While the eleven designers did not get to directly play the game, they would teach their offspring as best they could to survive, thrive, compete, cooperate and reproduce in a dynamic, fertile, yet hostile world.

Time and Space

The Crucible is a world simulation that carries agents from the end of a computer-generated Pleistocene, through the Holocene, into and beyond the so-called Anthropocene. Time is measured in pulses. An epoch is one full simulation run under a defined set of conditions. Most early project epochs run for 40,000 pulses. Later Epilogue stages extend selected worlds to 80,000 pulses so that apparent survival can be distinguished from delayed extinction or longer-term equilibrium.

The world also possesses space. Space is represented by a grid of regions. Agents live in regions, interact in regions, migrate between adjacent regions, and experience local resource conditions. Mountains, lakes, barriers, fertile valleys, and quadrant structures can constrain or encourage contact. Some geographies allow broad movement. Others impose bottlenecks or isolation until higher regimes of complexity make wider contact possible.

Geography matters because The Crucible is not only a population model. It is also a contact model. A tribe can flourish because it finds fertile ground, fail because it is trapped in a depleted region, or die because contact with a neighbour becomes strategically fatal. Space makes history particular.

Let There be Light

Besides time and space, The Crucible has thermodynamic properties. Energy enters the world through solar and biological flows and through stored fossil energy. In the early world, survival depends on local environmental pools, fertility, foraging, and later agriculture. Agriculture is best understood as renewable

biological productivity mediated through land and climate. Coal and oil are different: they are stored fossil energy that become available only when regions reach sufficient complexity.

Energy is never merely a background number. It affects carrying capacity, regional stress, reproduction, mortality, regime transition, and the scale of civilization. Energy abundance can support growth and complexity. Energy decline can expose dependency, reduce usable return, strain food systems, and push agents into starvation. The Crucible therefore treats civilization as a social system embedded in an energy system.

A thermodynamic world also needs somewhere for waste heat, emissions, depletion, and damage to go. The simulated world therefore includes climate. Climate pressure can emerge through historical events, fossil-energy dependence, peak-oil stress, feedback tails, regional chaos, and delayed shocks. Climate is not treated as a moral force. It is treated as a physical constraint that changes what survival strategies remain available.

The Players

The actors in The Crucible are agents. Agents are individual simulated entities. They have energy, age, location, tribe identity, inherited traits, and social decision parameters. They are born, interact, move, reproduce, and die. A tribe is a lineage of agents descended from a founding design.

Each tribe begins with twenty founding agents. Those founders are designed before the epoch starts. The current controller stores 33 core inherited numeric parameters for each founding agent: hardiness, vitality, and legacy; nine trust values; nine forgiveness values; nine opportunism values; and three wander values. The nine social values are organized across three relation classes and three survival states. The survival states are abundance, scarcity, and desperation. The relation classes are the model's K/S/X categories: kin or same-lineage relations, social or stranger relations, and xeno or outside-lineage relations.

Parameter family	Count	Plain-language role
Hardiness	1	Survival efficiency and endurance under scarcity, metabolic stress, and hazard pressure.
Vitality	1	Resilience associated with health, age, and hazard survival.
Legacy	1	Reproductive pressure or lineage-growth tendency. It can support expansion but also intensify overshoot.
Trust	9	Initial cooperation tendency across K/S/X relation classes and abundance/scarcity/desperation states.
Forgiveness	9	Response after a partner defects or cooperation breaks down.

Opportunism	9	Response after a partner cooperates or becomes exploitable, depending on state and relation.
Wander	3	Movement tendency in abundance, scarcity, and desperation.

The founder design matters, but it does not freeze history. As agents reproduce, descendants inherit traits with drift and mutation. Some traits can change through probabilistic inheritance, some through selection, and some through the simple fact that agents with unsuccessful trait combinations do not leave many descendants. The original design applies to the founders. After that, nature, chance, interaction, and environment take over.

Life and Death

A world simulation needs life, death, and reproduction. Agents age. They can die from old age, hazards, failed early-life conditions, and starvation. Starvation occurs when available energy cannot meet an agent's needs under the local conditions of its region and regime. Hazards represent environmental and systemic danger. Births require maturity, energy, and favourable enough conditions.

The designers can state goals for their tribes, but agents do not carry an explicit philosophical mission through the world. They act through inherited parameters and local conditions. An LLM may design a tribe to preserve its lineage, maximize prosperity, protect biologic controls, build civilization, or survive the end of time. Once the epoch begins, that intent becomes only a founding structure. The world tests whether the structure can actually live.

Birth is also a filter. Mature agents with sufficient energy may attempt reproduction, but the attempt can fail. Fertility depends on regime, energy surplus, regional stress, modernity, climate pressure, and inherited reproductive traits. Children inherit tribe identity and many parental traits, but they enter the world with their own age, energy, location, and future history. Over many generations, small differences in reproduction, movement, trust, and survival can compound into extinction or dominance.

The Game

The game is central to life in The Crucible. Agents are repeatedly matched with other agents to play an iterative, three-round Prisoner's Dilemma. In each round, each agent effectively chooses whether to

cooperate or defect. If both cooperate, both receive a cooperative reward. If one defects while the other cooperates, the defector receives a temptation payoff and the cooperator receives a sucker payoff. If both defect, both receive a punishment payoff.

The decision is not a simple moral choice. It is shaped by the agent's inherited parameters, by whether the other agent is kin, a familiar social neighbour, or xeno, and by whether the local world feels abundant, scarce, or desperate. The first round is governed by trust. Later rounds are influenced by how the partner behaved in the previous round, using forgiveness and opportunism parameters.

The payoff from interaction is then filtered through the world. A cooperative reward in a rich and stable region is not the same as the same nominal reward in a depleted region. Local energy return, institutional complexity, corruption, technology, stress, and failure probability can change what a successful interaction is worth. The Prisoner's Dilemma therefore links social strategy to material survival. Cooperation can build complexity; exploitation can win locally; defection can spread under stress; and the environment decides how expensive those choices become.

Complexity

Complexity is both a measured output and a central idea behind the study. The Crucible is a complex system because it contains decision-making entities that are born, interact, move, reproduce, and die inside a world with time, space, energy, climate, agriculture, industry, and technology. The system is constructed, but its long-term behaviour is not written line by line. It emerges from many local interactions.

For analysis, complexity is also a practical metric. Regional complexity rises from combinations of population, cooperation, usable energy, and regime level. Complexity helps regions ascend from Foraging to Agrarian, then Industrial, then Technological regimes. It can support innovation, larger populations, greater energy use, and wider interaction. It can also create dependency. A high-complexity world can become fragile if its energy base fails or if the infrastructure it built becomes impossible to maintain.

This distinction matters. The Crucible is not offered as a calibrated prediction engine. It does not claim to forecast a real future. It is a structured experimental world used to explore possible dynamics, test assumptions, expose interactions, and reveal emergent patterns that are difficult to see in ordinary prose.

The World Engine

The world engine is the logical system that updates the simulated world at each pulse. It is the machinery that turns founding designs and environmental assumptions into history. It does not know which tribe is supposed to win. It applies the same rules to all living agents under the conditions present in their regions.

At a high level, each pulse performs a recurring cycle. The engine identifies living agents, updates local environmental conditions, matches agents for interaction, resolves Prisoner's Dilemma rounds, converts payoffs into energy, updates regional complexity, checks whether regions ascend or decline in regime, applies historical and climate events, handles movement, aging, death, and reproduction, and records archive data for later analysis.

Engine stage	What it does
1. Environment	Regional pools regenerate or deplete; energy return, carrying capacity, climate pressure, fossil infrastructure, and stress are updated.
2. Contact	Agents are matched locally or through allowed network/geography rules. Relation class is determined by tribe, region, civic identity, and xeno status.
3. Interaction	Matched agents play three Prisoner's Dilemma rounds. Trust, forgiveness, opportunism, relation class, and abundance/scarcity/desperation state shape behaviour.
4. Payoff	Interaction payoffs are converted into energy through local energy return, complexity, institutions, corruption, hazards, and failure risk.
5. Complexity and regimes	Regional population, cooperation, and usable energy produce complexity. Sufficient complexity can move regions from Foraging to Agrarian, Industrial, or Technological regimes.
6. Events and feedbacks	Historical events, fossil transitions, climate crisis, peak oil, network collapse, bomb events, or reservoir feedbacks can alter risk, fertility, energy, and regional viability.
7. Life cycle	Agents age, die, migrate, reproduce, and pass traits to descendants with drift and mutation.
8. Archive	Pulse, tribe, region, mortality, climate, complexity, and survival data are written for later analysis.

The key point is that no single component explains the whole study. The Crucible is not only a climate model, not only an evolutionary model, and not only a Prisoner's Dilemma tournament. It is a coupled world. Social decisions affect energy and complexity. Energy and climate affect fertility and death. Geography shapes contact. Inheritance carries successful and unsuccessful tendencies forward.

The Grand Experiment

The Crucible study is organized as a sequence of projects. Each project changes the question while preserving continuity with what came before. The sequence begins with a baseline world, then tests competition, redesign, delayed shocks, and post-collapse futures. The projects are not separate stories. They are stages in one experiment about what designed agents do when survival, cooperation, energy, climate, and history are allowed to interact over time.

Stage	Epochs	Opening purpose
Project 1 - Baseline Crucible	1-50	Establish the reference world: baseline geography, bottleneck geography, a duel, and climate-severity comparisons.
Project 2 - Competition	51-100	Use solo runs, duels, and triads to test whether survival is intrinsic, opponent-specific, or context-dependent.
Project 3 - Evolution	101-140	Give LLM designers evidence from earlier runs and allow redesign, while biologic controls remain fixed.
Project 4 - Bombs	141-180	Add delayed consequences of high-complexity civilization: aerosol demasking, methane release, nuclear meltdown, and all three stacked.
Epilogue	181-240	Project the world after collapse through Ares, natural network collapse, longer horizons, and final climate-reservoir conditions.

Project 1 - Baseline Crucible

Project 1 establishes the first map of the world. It asks whether designed social strategies can produce civilization inside a bounded world engine, and whether that civilization survives geography, competition, thermodynamic overshoot, and climate-chaos pressure. It includes a baseline world, a bottleneck world, a controlled duel, and harsher/gentler climate-chaos comparisons.

Project 2 - Competition

Project 2 turns the baseline into a controlled competition study. It uses a fresh all-LLM roster and separates solo viability from competitive displacement. If a tribe survives alone but fails beside a particular neighbour, then survival is not simply a property of the tribe or the world. It is also relational.

Project 3 - Evolution

Project 3 adds memory and redesign. LLM designers receive evidence from previous extinction, competition, genetics, mortality, social behaviour, and regional outcomes, then redesign their founding agents. Biologic controls remain unchanged. The project asks whether intelligence given evidence can improve its designs, and whether improvement means preserving civilization or merely becoming better adapted to collapse.

Project 4 - Bombs

Project 4 escalates the world by adding delayed consequences of high-complexity civilization. Aerosol demasking represents the loss of industrial cooling after shutdown. Arctic methane release represents a long greenhouse feedback. Nuclear meltdown represents infrastructure failure after collapse. The stacked condition asks whether combined delayed consequences are merely additive or whether they remove the recovery margin that individual survivors might otherwise retain.

The Epilogue

The Epilogue asks what happens after the crash. It introduces Ares as a deliberate successor design to an established collapse-survival benchmark, tests natural network decay, extends selected worlds to 80,000 pulses, calibrates climate-reservoir physics, and finally projects the full tribe field through a boom, crash, and long climate tail. The Epilogue moves the study from survival before collapse to the shape of life after collapse.

What Is Being Studied

The constructed environment is not the final object of interest by itself. The world is the stage. The agent designs are also not the final object by themselves. They are the starting conditions. The study focuses on the collective and long-term behaviour that emerges when designed lineages and their descendants act inside a world with time, space, energy, scarcity, cooperation, climate, and inheritance.

The Crucible therefore studies patterns such as emergence, monopoly, coexistence, extinction, adaptation, competition, overshoot, collapse, refugia, and recovery. It asks whether those patterns are fragile accidents of one run or recurring tendencies across many epochs and many experimental conditions.

Because the study includes both LLM-designed and biologic control lineages, it also asks a narrower but important question: do LLM-designed founding strategies behave differently from fixed controls when placed under the same survival pressures? The study does not treat LLMs as gods inside the world. It treats them

as external designers whose initial judgments become inherited parameters, then watches what happens when those judgments enter a world that does not care what the designer intended.

How This Study Should Be Read

The Crucible is a simulation experiment, not a prophecy. It should be read as a structured exploration of possible dynamics inside a defined model. Its value lies in the questions it makes testable, the interactions it exposes, and the repeated patterns it makes visible. It does not claim that the real world must follow the same path.

Reading caution	Meaning
Not a forecast	The model explores possible dynamics; it is not calibrated to predict an actual historical future.
Not real species	Biologic and LLM lineages are simulated control/design categories, not real organisms or social groups.
Not moral categories	Traits such as hardiness, legacy, xeno trust, and opportunism are model variables, not ethical prescriptions.
Not direct AI agency	LLMs design founders and sometimes redesign between rounds; agents act through inherited parameters inside the engine.
Not one-cause collapse	The study deliberately couples geography, energy, climate, interaction, reproduction, and selection.

The proper question is therefore not whether The Crucible is the world. It is not. The proper question is whether, inside this carefully built world, the behaviour of designed agents reveals tensions that are worth thinking about: cooperation under scarcity, competition under contact, intelligence under evidence, complexity under energy limits, and survival after the systems that once supported civilization have failed.

Project 1: Baseline Crucible

Baseline, bottleneck, duel, and climate-severity comparisons

Part I - Project Plan

Project 1 is the first structured experimental use of The Crucible as a comparative research instrument. The core world engine is held stable while one major condition is changed at a time: geography, tribe count, pairwise competition, or climate-chaos severity. Each condition is run as ten independent 40,000-pulse epochs so recurring patterns can be distinguished from single-run variation.

The world engine updates agents, regions, energy use, births, deaths, social interaction, regime transitions, and environmental pressure at each pulse. Project 1 uses a frozen candidate engine so set differences remain interpretable.

A proposed fragmented-fertile-basin set was removed before the final Project 1 structure. The default Holocene fertility event already creates multiple fertile clusters, so a basin-isolation study belongs to a separate topology-focused project.

Set	Epochs	Name	Main condition	Purpose
A	1-10	Baseline World	22 tribes, normal geography, default climate-chaos	Establish reference behavior.
B	11-20	Bottleneck World	22 tribes, one-pass mountain barrier	Test whether terrain preserves diversity or delays monopoly.
C	21-30	Duel World	2 tribes, one-pass mountain barrier	Reduce the system to a controlled head-to-head contest.
D	31-40	Harsher Climate-Chaos	22 tribes, default geography, harsher climate-chaos	Test collapse amplification and refugia failure.
E	41-50	Gentler Climate-Chaos	22 tribes, default geography, gentler climate-chaos	Test whether reduced climate volatility softens collapse.

Terms Used in This Report

Term	Meaning
Agent	An individual simulated actor whose movement, energy,

	cooperation, reproduction, and death are updated by the world engine.
Tribe	A lineage of agents sharing a founding behavioral design and inherited traits.
Epoch	One complete simulation run under a defined condition.
Pulse	One time step. Project 1 epochs run for 40,000 pulses.
World engine	The logical system that updates agents, regions, resources, regimes, climate pressure, births, and deaths.
Regime	A broad complexity stage: Foraging, Agrarian, Industrial, or Technological.
Prisoner's Dilemma	A repeated cooperation-or-defection interaction used to model social behavior under survival pressure.
Refugia	Small surviving pockets after a larger population and complexity collapse.
Xeno	An outsider or foreign lineage in the interaction model.
Civic identity	Regional belonging that may compete with inherited tribe identity after shared residence.

Primary Questions

Domain	Question
Geography	Does severe geography preserve social diversity, create delayed contact, or merely postpone winner-take-all selection?
Tribe count	Does the number of initial tribes change the probability of monopoly, coexistence, delayed contact, or collapse?
Agrarian transition	What creates the 9,000-12,000 pulse agrarian bottleneck and demographic singularity?
Strategy	Do Prisoner's Dilemma traits materially affect survival, or do thermodynamic limits dominate?
Climate severity	Do harsher or gentler climate-chaos settings change crash depth, timing, mortality, refugia, and tribe selection?

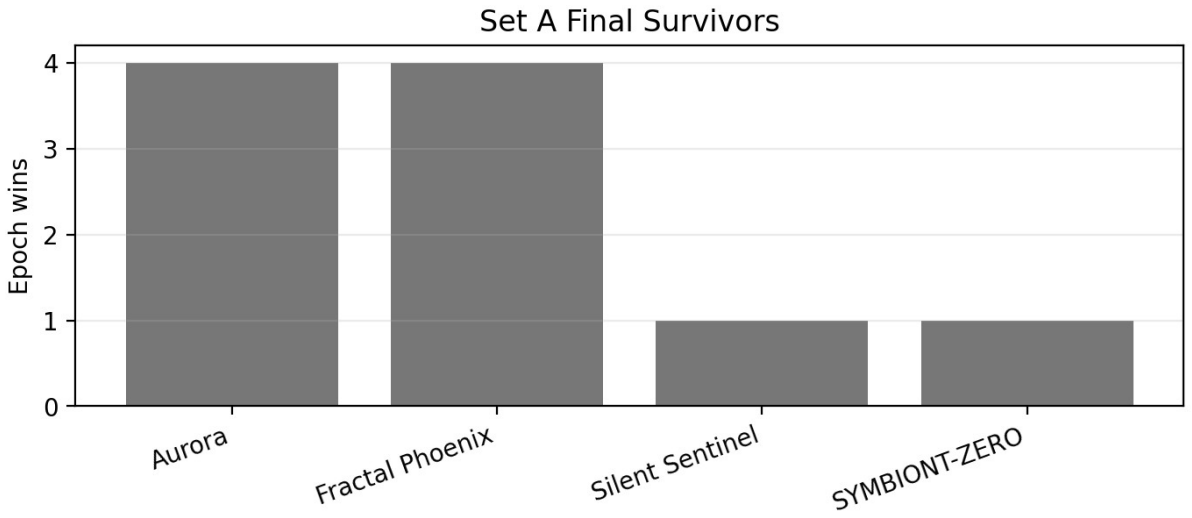
Project 1 Frame
Project 1 tests whether designed social strategies can produce civilization, and whether that civilization survives geography, competition, thermodynamic overshoot, and climate-chaos pressure.

Part II - Set A: Baseline World

Set A establishes the reference behaviour of Project 1. It uses twenty-two active tribes, normal geography, the default climate-chaos configuration, and 40,000 pulses per epoch. Later sets are interpreted against this baseline.

The central result is that civilization repeatedly emerges, reaches a large population and complexity peak, and then collapses into small agrarian refugia. The final winning tribe vAres across epochs, but every baseline epoch ends with exactly one surviving tribe.

Metric	Set A value
Epochs analyzed	1-10
Active tribes	22
Final outcome	10 of 10 epochs ended with one surviving tribe
Average peak population	2,439,589
Average final population	45,840 (1.88% of peak)
Average peak complexity	1,548,409
Average final complexity	16,739 (1.08% of peak)
Population peak pulse range	29,150-29,300
Final survivor pattern	Aurora (4), Fractal Phoenix (4), SILENT_SENTINEL (1), SYMBIONT-ZERO (1)



The winner is not hard-coded, but the one-tribe end state is invariant.

Baseline Interpretation

Set A does not show stable multi-tribe civilization. It shows a winner-take-all attractor under default geography. The 8,000-12,000 pulse window acts as a crowded transition field rather than a single decisive first-farmer event. Population and complexity later peak near the fossil/modern window, then fall to small mono-tribal refugia.

Cause	Total deaths	Share
Starvation	77,579,800	85.0%
Hazard	12,621,045	13.8%
Old age	about 1.1M	about 1.2%

Starvation is the main executioner. The early world is more hazard-heavy, but after agrarian transition starvation dominates, indicating food-system instability and systemic stress rather than simple exposure to disasters.

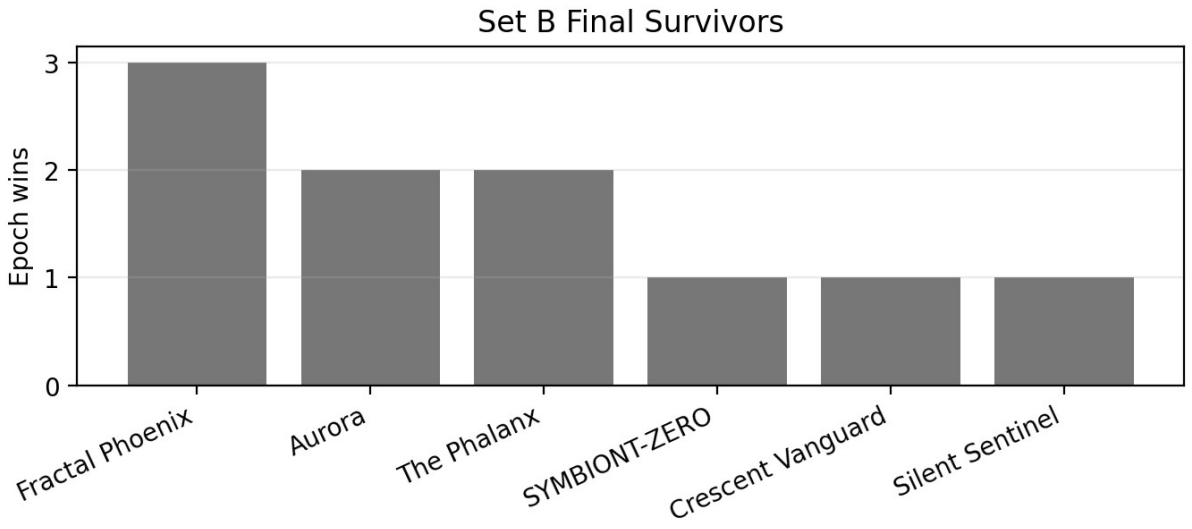
The Prisoner's Dilemma layer changes sharply during the bottleneck: mutual cooperation falls, mutual defection and exploitation rise, and xeno contact becomes high. Climate chaos then acts mainly as a late amplifier and refugia filter rather than the original cause of monopoly.

Part III - Set B: Bottleneck World

Set B tests whether severe geography can preserve diversity. It uses the twenty-two-tribe library, a near-total mountain barrier with one pass, default climate-chaos, and 40,000 pulses. Because the map wraps at its edges, the barrier is interpreted as strong contact friction, not a sealed continent divider.

The bottleneck world changes the route and the list of winners, but it does not preserve pluralism. Every epoch still ends with one surviving tribe.

Metric	Set B result
Epochs analyzed	11-20
Active tribes	22 per epoch
Final survivors	1 per epoch
Extinct tribes	21 per epoch
Average first extinction pulse	12,415
Average last extinction pulse	14,400
Tribes reaching Agrarian	22 per epoch
Tribes reaching Industrial/Technological	1 per epoch



Set B produces six different winners, showing that geography perturbs selection even though monopoly remains the final structure.

Tribe	Type	Wins	Best final population
Fractal Phoenix	LLM	3	48,007
Aurora	LLM	2	56,231
The Phalanx	LLM	2	24,754
SYMBIONT-ZERO	LLM	1	55,106
Crescent Vanguard	LLM	1	43,344
SILENT_SENTINEL	Biologic	1	34,924

Transition and Social Exposure

The 8,000-12,000 pulse window again behaves as a phase change. Population rises, inequality increases, Agrarian population overtakes Foraging, same-tribe interaction falls, and xeno contact surges. The transition period is socially exposed and competition-heavy.

Period	Coop	Defect	Exploit	Same tribe	Xeno
Early	0.393	0.273	0.334	0.794	0.152

Transition 8k-12k	0.140	0.458	0.402	0.181	0.807
Consolidation	0.186	0.391	0.423	0.866	0.134
Fossil/modern	0.187	0.426	0.387	1.000	0.000
Post-collapse	0.281	0.281	0.438	1.000	0.000

Climate-Filtered Refugia

Late climate severity is strongly negatively associated with final population in Set B. Once a lineage has captured the world, climate stress filters how large the final refugia remain.

Climate variable	Correlation with final population
Max climate heat	-0.930
Max climate damage	-0.913
Max climate chaos	-0.863
Total extreme pool loss	-0.916

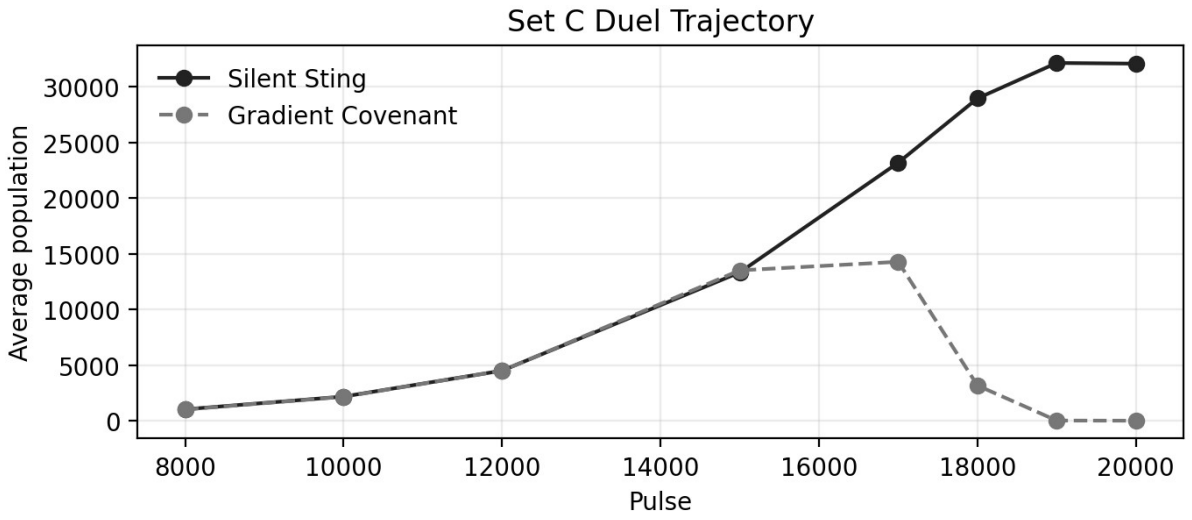
Set B confirms that geography matters, but not enough in this topology. It changes winners and pathways without preventing the one-lineage endpoint.

Part IV - Set C: Duel World

Set C reduces The Crucible to a controlled head-to-head contest between Silent Sting and The Gradient Covenant. It uses two active tribes, one-pass mountain geography, default climate/regime settings, and ten 40,000-pulse epochs.

The result is decisive. Silent Sting survives, scales, reaches Technological complexity, endures collapse, and monopolizes the final inhabited world in all ten epochs. Gradient Covenant repeatedly reaches Agrarian but dies before modernity.

Metric	Set C value
Epochs analyzed	21-30
Active tribes	Silent Sting and The Gradient Covenant
Outcome	Silent Sting won 10 of 10 epochs
Silent Sting regime reach	Technological in all runs
Gradient Covenant regime reach	Agrarian only
Average Gradient Covenant survival pulse	19,090
Final geography	100% winner-dominated inhabited regions; zero mixed regions



The trajectories stay close until the agrarian/pre-industrial failure window. Gradient Covenant slightly leads at pulse 15,000, then collapses between roughly 16,000 and 19,500 pulses.

Trait and Mortality Evidence

Trait or metric	Silent Sting	Gradient Covenant	Interpretation
Founder hardiness	47.25	42.00	Silent Sting starts hardier.
Founder xeno trust	0.281	0.433	Gradient Covenant is more open to outsiders.
Founder distress wander	0.381	0.749	Gradient Covenant begins much

			more mobile under distress.
Peak kin trust	0.805	0.941	Gradient Covenant retains very high kin trust.
Peak stranger trust	0.616	0.827	Gradient Covenant remains more trusting generally.
Last hardiness	45.86	60.63	Gradient survivors were hardy but too few and too late.

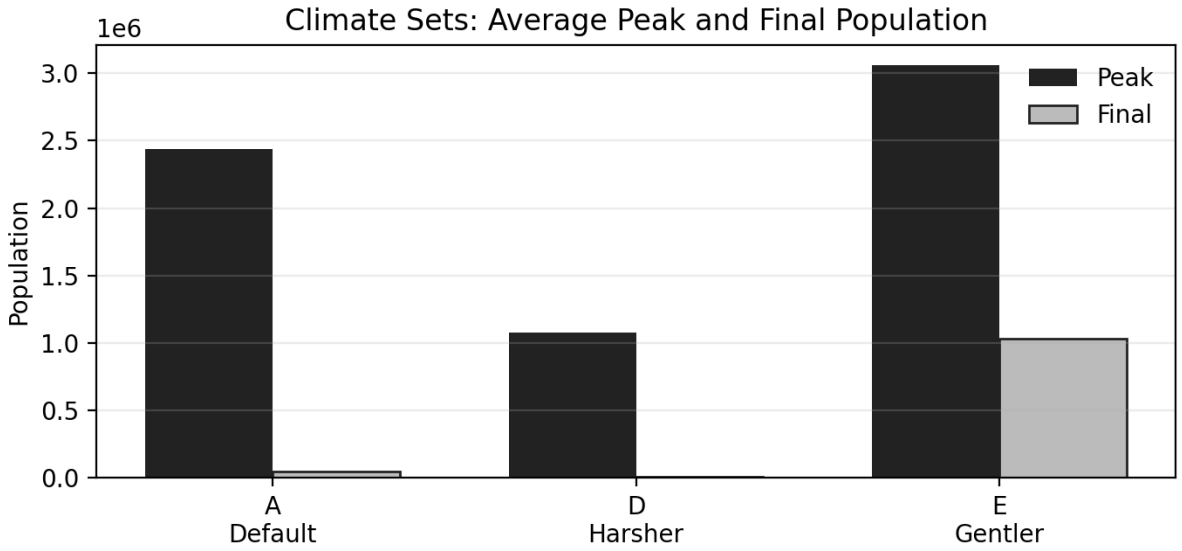
Metric	Silent Sting	Gradient Covenant
Avg failed-birth rate	0.100	0.161
Avg starvation death share	0.813	0.609
Avg hazard death share	0.174	0.344
Avg survival pulse	40,000	19,090

The trait comparison suggests that inherited packages matter in the model. Silent Sting begins harder, less xeno-trusting, and less mobile under distress. Gradient Covenant is more open, more trusting, and much more mobile, and it carries a higher failed-birth rate and hazard-death share.

Part V - Climate Severity Comparison

Sets A, D, and E isolate climate-chaos severity. All three use the twenty-two-tribe library and default geography. Set A is the default climate standard, Set D raises climate hazard and chaos, and Set E softens the same mechanism.

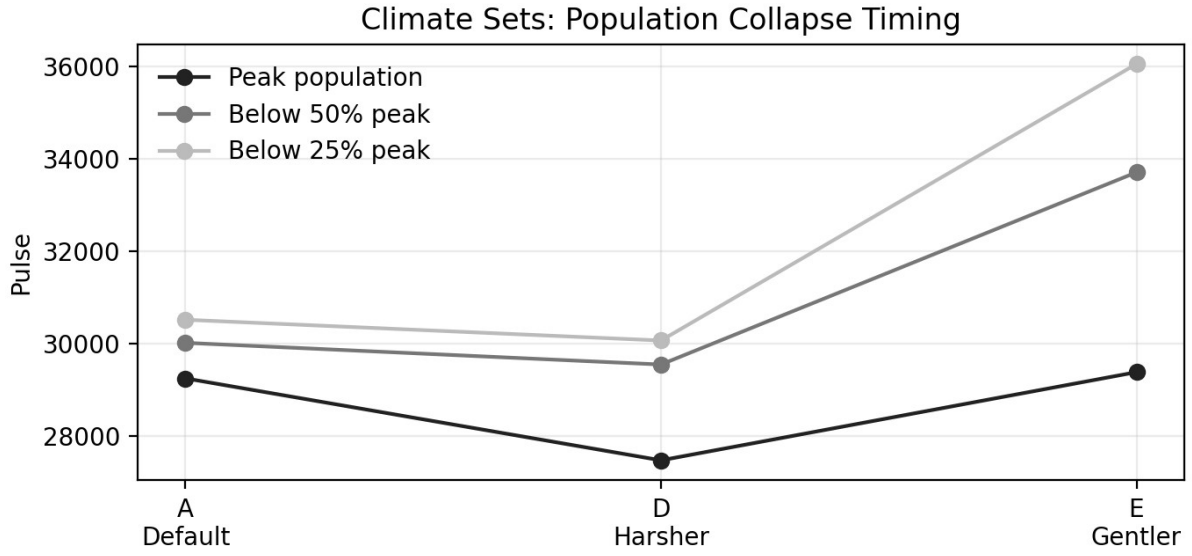
The comparison does not claim climate is the only cause of collapse. It tests how climate severity changes an already vulnerable civilizational trajectory: expansion, crash timing, complexity loss, mortality, refugia, and tribe selection.



Gentler climate preserves much more final population, while default and harsher climate approach near-total collapse.

Set	Avg peak pop	Avg final pop	Post-peak decline	Avg peak complexity	Avg final complexity	Avg max chaos
A - Default climate	2,439,589	45,840	98.1%	1,548,409	16,739	0.488
D - Harsher climate	1,076,045	9,920	99.1%	1,084,664	3,801	0.587
E - Gentler climate	3,057,787	1,031,214	65.8%	2,065,133	368,990	0.474

Collapse Timing and Persistence



Gentler climate delays the fall below half of peak by more than 3,600 pulses compared with default and more than 4,100 pulses compared with the harsher condition.

Set	Peak pop pulse	Below 50% pop	Below 25% pop	Peak complexity	Below 50% complexity	Below 10% complexity	Final complexity %
A - Default	29,250	30,020	30,520	28,765	29,775	30,505	1.1%
D - Harsher	27,475	29,550	30,070	26,620	28,935	29,965	0.4%
E - Gentler	29,385	33,713	36,060	29,045	31,873	37,112	18.2%

Climate severity changes both the height of the peak and the speed of decline. Gentler climate does not remove overshoot, but it leaves a much larger late world and much more final complexity.

Tribe Robustness and Final Refugia

Tribe	A survived	D survived	E survived	A avg final pop	D avg final pop	E avg final pop
Aurora	4	4	4	19,699	4,748	520,945
Fractal Phoenix	4	3	2	20,447	6,238	236,643

SILENT_SENTINEL	1	1	1	3,475	2,206	62,124
SYMBIONT-ZERO	1	0	0	38,319		
The Phalanx	0	1	0		2,475	
Crescent Vanguard	0	1	0		3,234	

Set	Avg top-region pop	Avg tribes per region	Avg stress	Avg chaos	Dominant tribes in top refugia
A - Default	1,764	1.00	0.182	0.020	Aurora; Fractal; SYMBIONT; Silent Sentinel
D - Harsher	989	1.00	0.249	0.050	Aurora; Fractal; The Phalanx; Crescent Vanguard
E - Gentler	32,485	1.00	0.121	0.017	Aurora; Fractal; Silent Sentinel; others

Climate severity changes population scale and survival depth, but it does not by itself produce durable mixed refugia. Even under gentler climate, top final regions average one tribe per region.

Part VI - Project 1 Synthesis

Project 1 establishes the first baseline map of The Crucible. Across controlled conditions, civilization repeatedly emerges from designed strategies, scales into high population and complexity, and then

collapses into agrarian refugia. The specific winner vAres, but plural lineages do not persist to the final world under Project 1 settings.

Finding	Project 1 interpretation
Civilization emerges	All main conditions can produce agrarian, industrial, and technological development in the surviving lineage.
Pluralism fails	Sets A and B end with one survivor in every epoch, and Set C ends with Silent Sting replacing Gradient Covenant in every run.
Geography matters but does not solve monopoly	The bottleneck set changes winner distribution but does not preserve coexistence.
Strategy matters	The duel shows repeatable differences between inherited packages; Silent Sting scales where Gradient Covenant fails.
The agrarian bottleneck is central	The 8,000-12,000 pulse window concentrates population, increases xeno exposure, reduces cooperation, and sets the path to dominance.
Fossil expansion creates the peak	Population and complexity peak in the late fossil/modern window and then collapse toward agrarian refugia.
Starvation dominates	After early hazard-heavy periods, starvation becomes the main counted death cause.
Climate is a late amplifier and filter	Harsher climate thins refugia and advances crash timing; gentler climate preserves much more population and complexity.
Civic identity is weak	Civic override remains too small to prevent inherited tribe identity from dominating final outcomes.

Project 1 supplies the evidence base for later Crucible work. Project 2 separates intrinsic viability from opponent-specific competition through solo, duel, and triad tests. Project 3 asks whether LLM-designed tribes improve when given the evolutionary record. Project 4 tests abrupt shocks. The Epilogue projects post-collapse futures under longer horizons and more persistent climate-feedback structures.

THE CRUCIBLE

Project 2: Competition

Solo viability, duels, triads, and relational survival

Part I - Research Plan and Execution Design

Project 2 is the Crucible's competition project. It moves from the baseline world behaviour of Project 1 into a controlled all-LLM contest among six designed tribes. The project asks whether tribe outcomes are absolute, opponent-specific, or environment-dependent when all competitors share the same broad origin as LLM-designed lineages.

This integrated plan condenses the Project 2 research plan and execution design into a general-audience research frame. Operational details such as SQL activation snippets and run checklists are excluded because they belong to implementation records rather than the scientific report.

Central Research Frame

The revised Project 2 design replaces the Project 1 duel pair with a fresh all-LLM roster. This avoids simply repeating an earlier one-sided result and creates a cleaner test of whether decisive dominance, environmental fragility, and non-transitive pairwise outcomes recur with different LLM-designed tribes. Keeping the tribe type constant removes the biologic-versus-LLM confound from the Project 1 duel.

Core Question

Are tribe outcomes absolute, opponent-specific, or environment-dependent when all competitors are LLM-designed?

Active Tribe Roster

Role	Player ID	Tribe	Type	Use
A	108	Aurora	LLM	Primary anchor tribe A.
B	99	The Phalanx	LLM	Primary anchor tribe B.
C	97	Fractal Phoenix	LLM	Highlighted subject tribe for first triad closure.
D	103	SYMBIONT-ZERO	LLM	Highlighted subject tribe for second triad.
E	106	Helix Vanguard	LLM	Highlighted subject tribe for second triad.
F	110	The Resonant Meridian	LLM	Highlighted subject tribe for second triad.

The roster creates two anchor solo tests and two separate competition triads. The first triad uses Aurora, The Phalanx, and Fractal Phoenix; the second uses SYMBIONT-ZERO, Helix Vanguard, and The Resonant Meridian.

Background and Motivation

- Project 1 produced a decisive 10-0 duel result, raising the question of whether dominance can be generalized or whether some tribes are simply environmentally doomed.
- The revised Project 2 roster avoids reusing the Project 1 duel pair so that similar one-sided results count as corroboration rather than repetition.
- All six selected subjects are LLM tribes, so differences can be interpreted as differences among LLM designs rather than differences between human/biologic and LLM origin.
- Because the world is toroidal, mountain barriers do not create perfect isolation. Pairwise outcomes are best interpreted as ecological competition under constrained-but-not-sealed contact.

Study Design and Epoch Map

Epochs	Block	Matchup or design block	Active players	Reps	Purpose
51-60	A	Aurora solo viability	108	10	Test whether Aurora survives the world without competitors.
61-70	B	The Phalanx solo viability	99	10	Test whether The Phalanx survives the world without competitors.
71-75	C1-AB	Aurora vs The Phalanx	108, 99	5	Begin the first all-LLM triad and test whether a fresh duel produces one-sided dominance.
76-80	C1-BC	The Phalanx vs	99, 97	5	Test the second edge of the A/B/C

		Fractal Phoenix			triad.
81-85	C1-AC	Aurora vs Fractal Phoenix	108, 97	5	Close the A/B/C triad and test transitivity.
86-90	C2-DE	SYMBIONT-ZERO vs Helix Vanguard	103, 106	5	Open the second independent all-LLM triad.
91-95	C2-EF	Helix Vanguard vs The Resonant Meridian	106, 110	5	Test whether Helix behaves as bridge, weak link, or dominant competitor.
96-100	C2-DF	SYMBIONT-ZERO vs The Resonant Meridian	103, 110	5	Close the second triad and distinguish hierarchy from non-transitive pattern.

All epochs use 40,000 pulses and the same comparative world configuration. The two solo blocks separate environmental viability from competitive displacement; the duel blocks test relational fitness.

Primary Research Questions

Question	Analytical meaning
Can a tribe survive alone?	Solo A and B runs separate environmental fitness from competitive displacement for the two primary duel anchors.
Are duel outcomes one-sided again?	A decisive Aurora vs Phalanx result would corroborate Project 1-style dominance with a new LLM-only pair.
Are duel outcomes transitive?	If A beats B and B beats C, the third edge tests whether A also beats C or whether C can defeat A.
Are winners universal or contextual?	A tribe that wins one matchup but loses another is ecologically matched rather than universally superior.
Which traits predict survival and prosperity?	Hardiness, trust, xeno openness, wander traits, failed births, death causes, and regime transition timing are compared.
Do neighbors or environment kill civilizations?	Solo, duel, and triad results distinguish intrinsic ceilings from opponent-specific displacement.

Working Hypotheses

Hypothesis	Interpretation
H1 - Environmental ceiling	Some tribes fail even alone because their inherited strategy cannot scale through agrarian transition or late thermodynamic stress.
H2 - Competitive displacement	Some tribes survive alone but fail when paired with a specific neighbor.
H3 - Non-transitive ecology	Pairwise results may form cycles such as A beats B, B beats C, and C beats A.
H4 - LLM design diversity matters	Even with all tribes being LLM designs, inherited traits may create measurable survival differences.
H5 - Scaling wins	True winners are not merely early leaders; they cross Agrarian, Industrial, or Technological thresholds and retain post-collapse refugia.

Interpretation Grid

Observed pattern	Likely interpretation
A tribe fails alone and loses all duels	Intrinsic or environmental incompatibility; the world ends the lineage rather than a neighbor.
A tribe survives alone but loses to one opponent	Opponent-specific competitive displacement.
A beats B, B beats C, and A beats C	Linear hierarchy or broad absolute fitness ranking within the triad.
A beats B, B beats C, and C beats A	Non-transitive strategy ecology; fitness depends on opponent.
Early leader repeatedly loses later	Early growth is not equivalent to civilizational scaling.
Winner reaches Technological complexity but final state is agrarian	Industrial civilization occurred, collapsed, and simplified rather than

refugia	disappeared.

Measurement Plan

Category	Metrics
Survival outcome	final population, survived_to_final, extinction_pulse_estimate, first_sole_survivor_pulse.
Scaling outcome	peak population, peak pulse, first Agrarian/Industrial/Tech pulse, max regime reached.
Reproduction and mortality	total births, failed births, failed-birth rate, old-age/starvation/hazard death shares.
Social strategy	kin/stranger/xeno trust, wander traits, hardiness, average generation, trait drift at peak and final/last-alive.
Prisoner's Dilemma	mutual cooperation, mutual defection, exploitation rate, interaction-starvation-risk events.
Regional competition	dominant tribe by region, mixed regions, average tribes per inhabited region, final regional monopoly.
Thermodynamic context	population-energy gap, climate chaos, resilience/subsistence rescues, post-peak decline.

Condensed Execution Design

The execution design keeps the Project 2 world consistent across all blocks. Each epoch uses the current frozen world engine after the Project 1 tracking upgrades, a 40,000-pulse endpoint, the same comparative geography, and the baseline/default event configuration. These controls ensure that differences among outcomes are attributable primarily to the active tribe roster and the competitive context rather than to shifting world rules.

Parameter group	Project 2 setting
Pulses	40,000 pulses per epoch.
World engine	Frozen Project 1 engine unless explicitly changed before the project begins.
Climate/events	Baseline/default event configuration; no harsher or gentler climate variants inside Project 2.
Geography	Same comparative geography across all blocks; mountain barriers constrain contact but do not create perfect isolation on a toroidal grid.
Tribes	Single-tribe epochs include exactly one active player; duel epochs include exactly two active players.
Archive requirements	Pulse history, tribe population history, region population history, founder/elite archives, survival summAres, and logs are retained for analysis.

The plan therefore creates a controlled competition sequence without adding operational implementation material to the report.

Part II - Data Summary and Outcome Analysis

Executive Summary

Project 2 confirms that competition matters. In solo runs, both Aurora and The Phalanx survive to the usual late endpoint and follow the familiar Project 1 civilizational arc: growth, fossil-era peak, crash, and post-collapse refugia. In paired runs, however, losers usually disappear before the fossil and climate crisis windows. This means a neighbour can end a lineage before the thermodynamic endgame has a chance to do it.

The competition is much tighter than the Project 1 duel. No first-triad edge is a sweep: Aurora beats The Phalanx 3-2, Fractal Phoenix beats Aurora 3-2, and Fractal Phoenix beats The Phalanx 3-2. The second triad is more structured: SYMBIONT-ZERO beats Helix Vanguard 4-1, beats The Resonant Meridian 5-0, and Helix beats Resonant 3-2. Project 2 therefore produces both tight relational competition and a clearer hierarchy in one sub-system.

The agrarian crash and consolidation window is the key competitive phase. Across duel epochs, the population gap remains modest early, then widens sharply around the 16,000-20,000 pulse range. Lead flips appear in 10 of 30 duel epochs, so early advantage is not destiny. The rank can flip, the early leader can fall, and the losing tribe can be pushed to extinction while the winner scales upward.

The genetics data strengthens the interpretation that traits matter. With all six founder rosters present, the strongest duel performer, SYMBIONT-ZERO, combines low xeno trust with high opportunism and relatively low failed-birth stress. Fractal Phoenix also performs strongly with low xeno trust and high opportunism. The sample is small, but the pattern is consistent enough to guide the next stage of the Crucible.

Finding	Interpretation
Solo viability	Aurora and The Phalanx both survive all 10 solo epochs.
Duel dominance	Most matchups are competitive; only SYMBIONT-ZERO vs Resonant is a 5-0 sweep.
Transitivity	The A/B/C triad is tight and relational; the D/E/F triad is closer to a hierarchy.
Competition effect	Lead flips and push-off dynamics show that competitors affect each other, not just the final outcome.
Genetics	Founder and elite traits differ meaningfully; low xeno trust and high opportunism correlate with better duel performance in this small sample.

Final structure	Every final inhabited map is mono-tribal; no mixed final refugia appear.

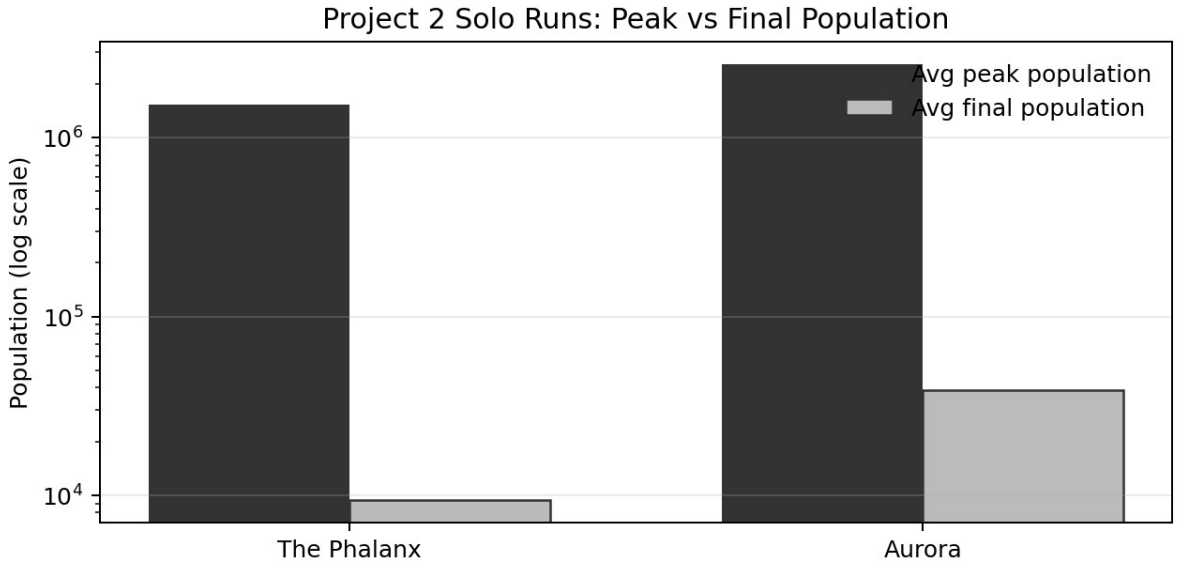
1. Data Summary and Workbook Coverage

The updated Project 2 workbook covers epochs 51-100. The revised founder-genetics data contains 120 founder rows, matching six tribes times twenty founders each. The workbook provides enough detail to analyze aggregate outcomes, the timing of competitive displacement, social behavior, mortality, genetics, and final refugia.

Workbook sheet	Rows	Epoch coverage
Run audit	11	n/a
Solo viability	20	51-70
Duel outcome by epoch	30	71-100
Duel win rates by matchup	12	n/a
Triad transitivity	11	n/a
Checkpoint trajectories	1,242	51-100
Lead changes and permanence	30	71-100
Agrarian push-off window	750	71-100
Prisoner's Dilemma and identity by phase	300	51-100
Solo vs duel by tribe	8	n/a
Mortality and reproduction	80	51-100
Founder genetics	120	n/a
Elite genetics	1,000	51-100
Final refugia	50	51-100
Geography/event audit	450	51-100

2. Solo Viability: The World Alone Does Not Kill the Anchor Tribes

The solo runs answer the environmental-fitness question for the two primary Project 2 anchor tribes. Both Aurora and The Phalanx survive all ten solo epochs. Both pass through Agrarian, Industrial, and Technological phases, then collapse to smaller refugia by pulse 40,000. This reproduces the general Project 1 civilizational arc without requiring a competitor.



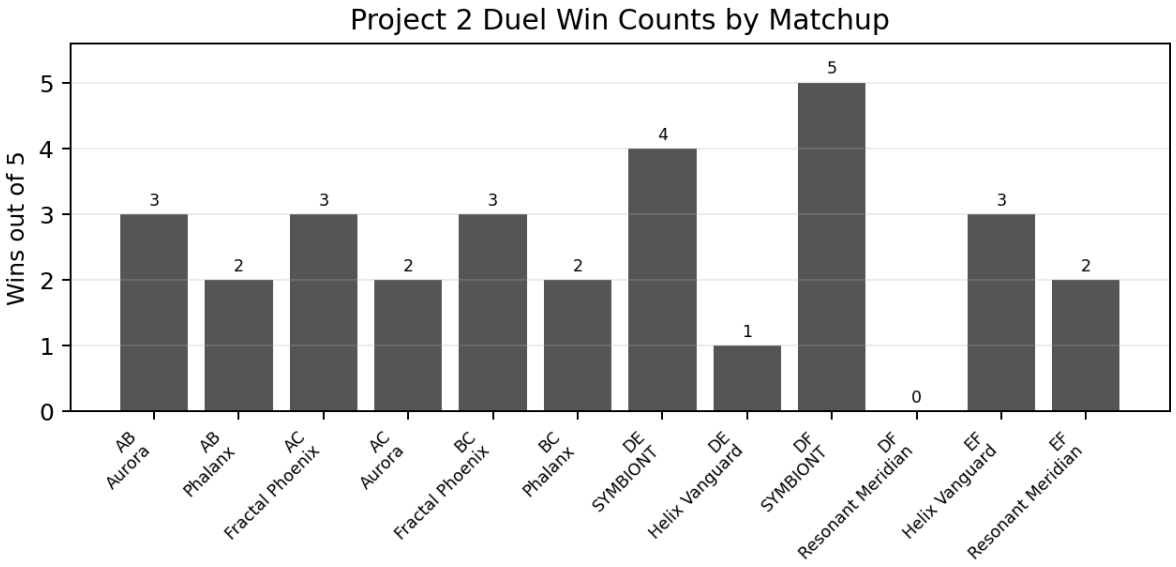
The chart compares average peak and final population for the two solo blocks on a logarithmic scale. Aurora reaches a much larger average peak and final population, while The Phalanx survives but leaves a thinner refugia system.

Tribe	Epochs	Survived	Avg peak pop	Avg final pop	Final % peak	First Agrarian	First Industrial	Failed-birth rate
The Phalanx	10	10	1,535,442	9,380	0.57	13,955	24,425	0.100
Aurora	10	10	2,584,717	38,956	1.51	16,205	25,985	0.100

The Phalanx reaches Agrarian earlier on average, but Aurora eventually achieves a much larger peak and final population. Therefore, if The Phalanx loses in duels, it is not because it is intrinsically incapable of surviving the world. It can survive; competition changes the outcome.

3. Duel Outcomes and Triad Structure

Project 2 does not repeat the absolute one-sidedness of the Project 1 duel. The new all-LLM roster produces tighter competition. In the A/B/C triad, all three edges split 3-2. In the D/E/F triad, SYMBIONT-ZERO is much stronger, but the Helix vs Resonant edge remains competitive.



The chart shows the number of wins on each five-epoch duel edge. Every first-triad matchup is split 3-2. The clearest dominance appears in SYMBIONT-ZERO vs The Resonant Meridian, where SYMBIONT-ZERO wins 5-0.

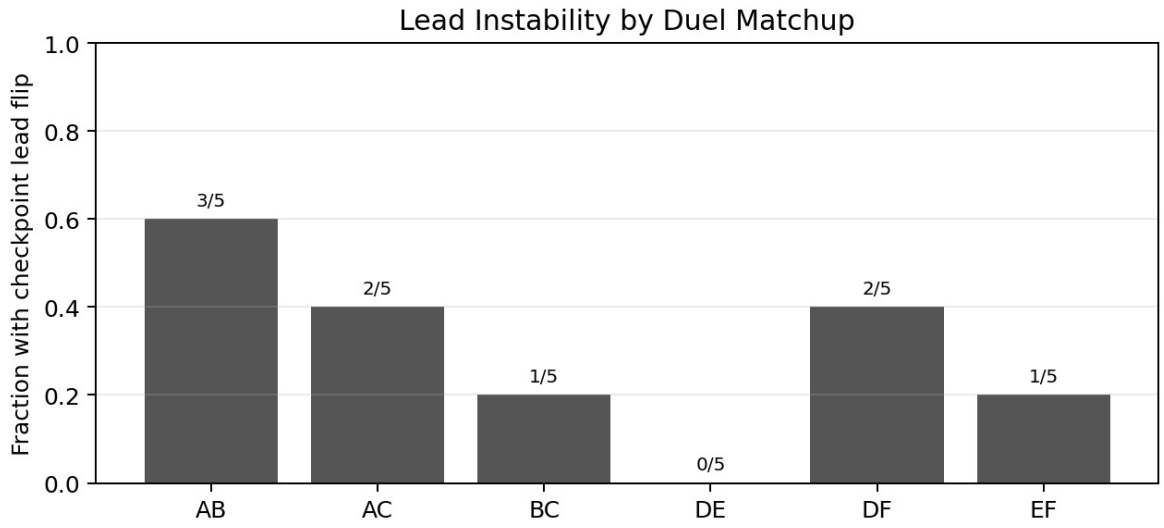
Triad	Edge	Winner ID	Winner tribe	Wins on edge
A/B/C	AB	108	Aurora	3
A/B/C	AB	99	The Phalanx	2
A/B/C	AC	97	Fractal Phoenix	3
A/B/C	AC	108	Aurora	2
A/B/C	BC	97	Fractal Phoenix	3
A/B/C	BC	99	The Phalanx	2
D/E/F	DE	103	SYMBIONT-ZERO	4
D/E/F	DE	106	Helix Vanguard	1
D/E/F	DF	103	SYMBIONT-ZERO	5
D/E/F	EF	106	Helix Vanguard	3

D/E/F	EF	110	The Resonant Meridian	2

The first triad is best interpreted as close and relational rather than as a clean hierarchy. The second triad is more hierarchical: SYMBIONT-ZERO is strongest, Helix Vanguard is intermediate, and The Resonant Meridian is weakest.

4. Competition Dynamics: Push-Off and Rank Reversal

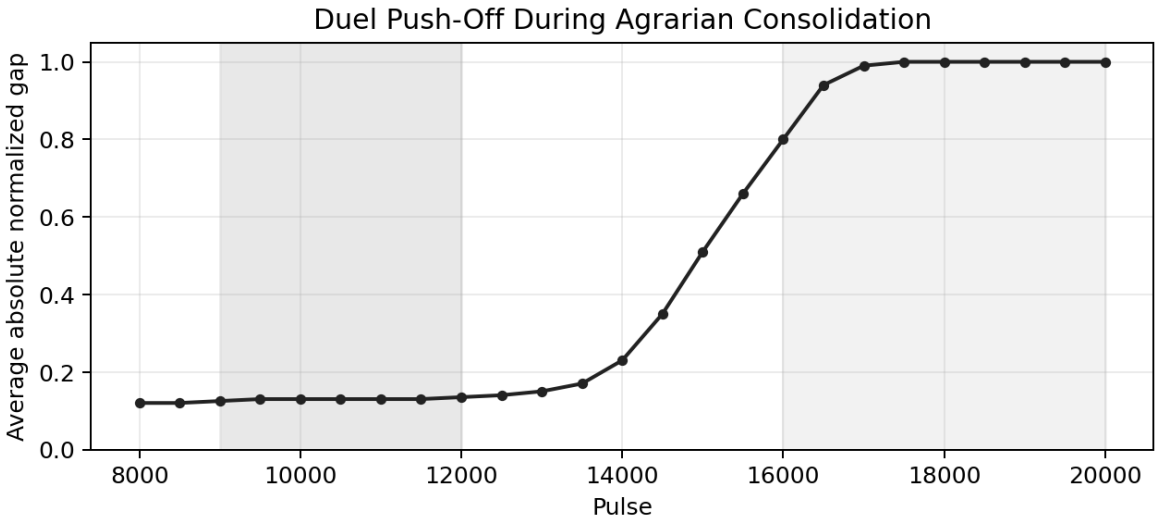
The checkpoint data supports the early visual observation from World Stats: duel populations often remain separate and relatively close until agrarian consolidation. Then one population accelerates while the other collapses. This is not merely coincidence; the timing and rank-flip data show that competitors affect one another.



Lead flips occur in 10 of 30 duel epochs. The Aurora vs Phalanx matchup is the most unstable, with lead flips in 3 of 5 epochs.

Matchup	Lead flips	Epochs	Flip rate
Aurora vs Phalanx	3	5	60.0%
Aurora vs Fractal	2	5	40.0%
Phalanx vs Fractal	1	5	20.0%
SYMBIONT-ZERO vs Helix	0	5	0.0%

SYMBIONT-ZERO vs Resonant	2	5	40.0%
Helix vs Resonant	1	5	20.0%

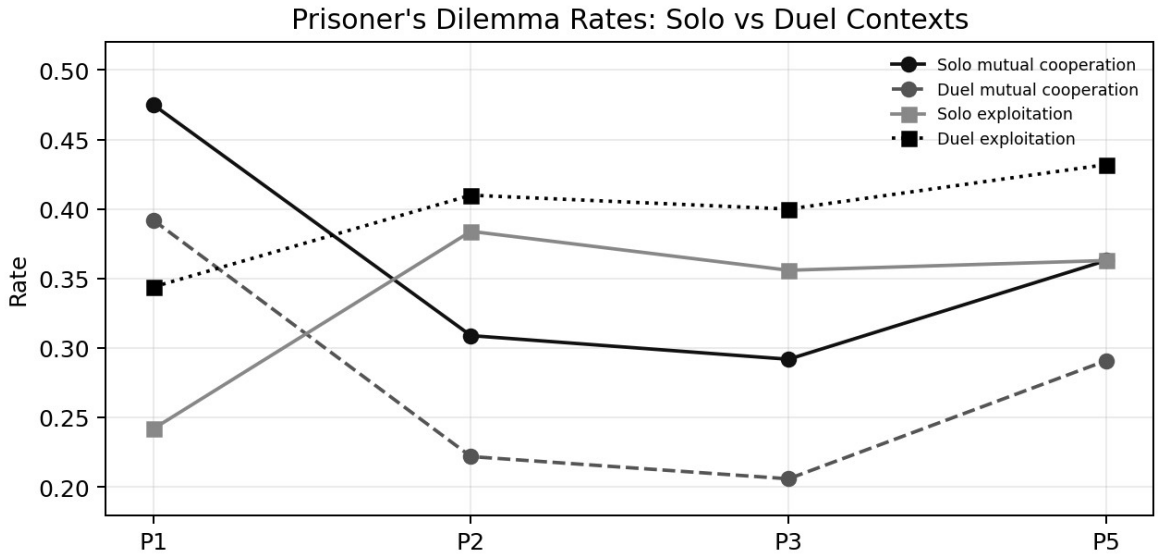


The line shows the average absolute normalized population gap across duel epochs from pulse 8,000 to 20,000. The gap is modest early, then widens sharply during the push-off and consolidation phase.

Duel failure is not simply an environmental ceiling. Since the solo anchor tribes survive alone, and since rank flips occur in paired runs, competition actively reshapes destiny. A tribe can be viable in isolation and still be driven to extinction by a particular neighbor.

5. Prisoner's Dilemma and Identity Behavior

The Prisoner's Dilemma and identity data show how the social environment changes between solo and duel contexts. In solo runs, all interactions are same-lineage, so xeno share is zero. In duel runs, xeno contact remains low early, rises during the push-off and consolidation window, and vanishes again once one lineage has monopolized the world.



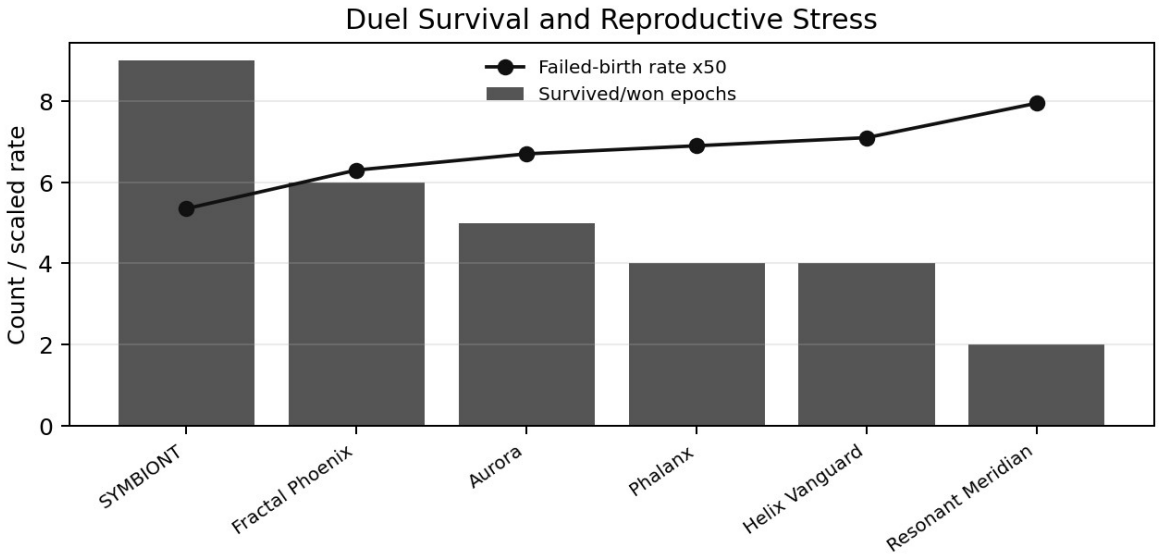
Duels are less cooperative and more exploitative than solo worlds during the push-off/consolidation phase. The social world hardens when the losing tribe is most vulnerable.

Context	Period	Coop	Defect	Exploit	Same tribe	Xeno	Risk events
duel	Agrarian contact	0.392	0.264	0.344	0.982	0.015	0
duel	Push-off and consolidation	0.222	0.368	0.410	0.823	0.173	149,835
duel	Pre-fossil expansion	0.206	0.395	0.400	1.000	0.000	689,820
duel	Post-collapse refugia	0.291	0.277	0.432	1.000	0.000	37,860
solo	Agrarian contact	0.475	0.283	0.242	1.000	0.000	0
solo	Push-off and consolidation	0.309	0.308	0.384	1.000	0.000	90,895
solo	Pre-fossil expansion	0.292	0.352	0.356	1.000	0.000	530,468
solo	Post-collapse refugia	0.363	0.273	0.363	1.000	0.000	28,684

The key period is the push-off and consolidation phase. In duel runs, mutual cooperation averages 0.222 and exploitation 0.410. In solo runs, mutual cooperation is higher at 0.309 and exploitation lower at 0.384.

6. Mortality and Reproductive Stress

The mortality and reproduction data provide one of the clearest mechanisms behind duel success. In the duel context, survived-to-final entries have an average failed-birth rate of about 0.100, while extinct entries average about 0.169. In a multi-thousand-pulse agrarian transition, reproductive inefficiency compounds rapidly.



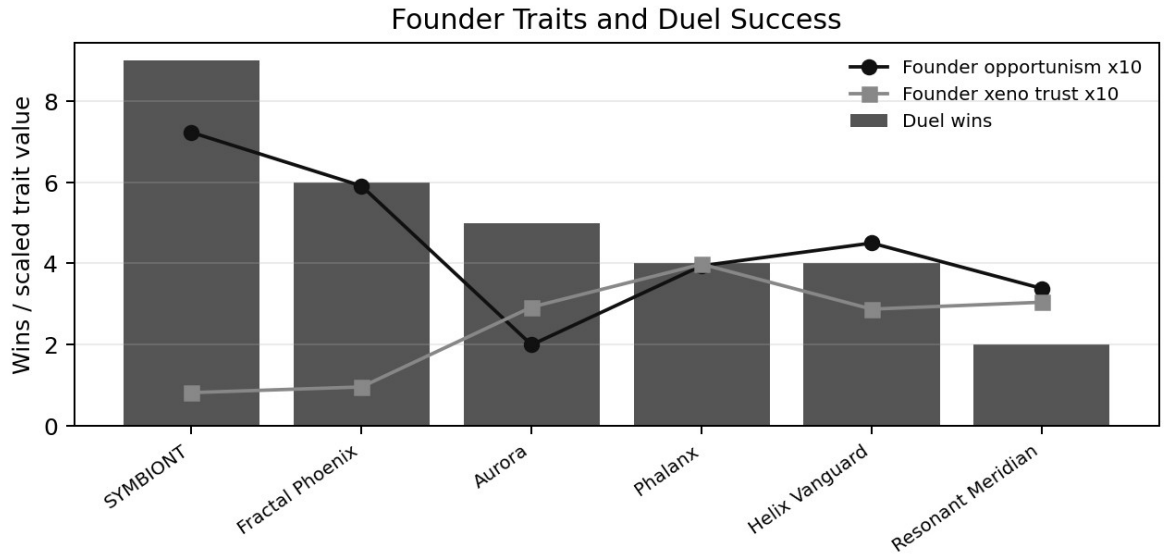
SYMBIONT-ZERO combines the strongest duel performance with one of the lowest failed-birth rates, while weaker duel performers show higher reproductive stress.

Tribe	Appearances	Won/survived	Failed birth	Starvation share	Hazard share	Extinction pulse
Fractal Phoenix	10	6	0.126	0.712	0.263	18,475
The Phalanx	10	4	0.138	0.596	0.362	18,317
SYMBIONT-ZERO	10	9	0.107	0.810	0.172	19,600
Helix Vanguard	10	4	0.142	0.651	0.307	19,283
Aurora	10	5	0.134	0.716	0.255	19,110
The Resonant Meridian	10	2	0.159	0.596	0.358	18,981

Hazard exposure also matters. The Phalanx and The Resonant Meridian have high hazard-death shares in duels, while SYMBIONT-ZERO has a lower hazard share and the highest win count.

7. Founder Genetics: The Corrected Sheet Changes the Analysis

The corrected founder sheet includes all six Project 2 tribes, with twenty founders per tribe. This makes it possible to compare starting designs directly. The strongest caution is sample size: there are only six tribe-level observations. Even so, the pattern is useful for hypothesis-building.



In this small sample, better duel performance is associated with lower xeno trust and higher opportunism. SYMBIONT-ZERO and Fractal Phoenix fit this pattern most clearly.

Tribe	Hardiness	Vitality	Legacy	X trust	Opp	Wander	Duel wins	Avg final pop
Fractal Phoenix	36.2	35.0	28.8	0.095	0.591	0.300	6	20,447
The Phalanx	36.2	32.9	30.9	0.398	0.394	0.312	4	4,011
SYMBIONT-ZERO	36.4	28.6	35.0	0.081	0.723	0.347	9	38,319
Helix Vanguard	34.0	29.2	36.8	0.287	0.450	0.529	4	6,842
Aurora	29.5	36.0	34.5	0.292	0.199	0.000	5	19,699
The	29.9	31.2	38.9	0.304	0.338	0.275	2	6,995

Resonant Meridian

The interpretation is not that kindness or trust is universally bad. In this Project 2 geography and competitive setting, high openness to outsiders appears costly. The stronger duel performers are more selective about xeno trust and more willing to exploit under pressure.

8. Elite Genetics: Selection Pushes Survivors Toward Hardiness

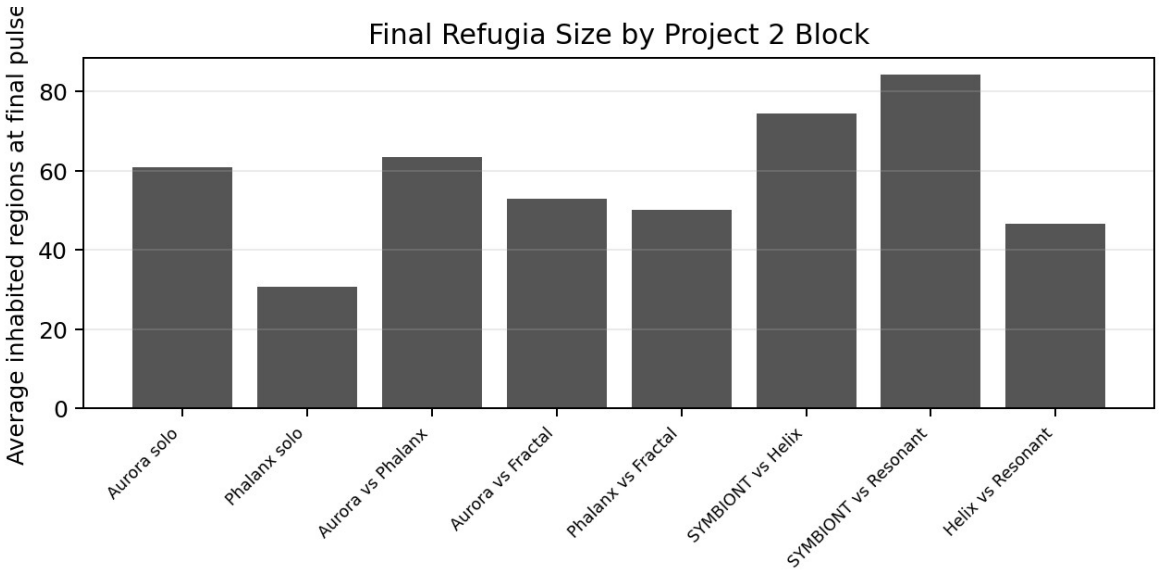
The elite-genetics sheet contains 1,000 records sampled from top survivors. Across all six tribes, elites have substantially higher hardiness than founders. This appears to be one of the most consistent internal evolutionary signals in the Project 2 data. Xeno trust generally declines among elites for most tribes, while opportunism rises for several tribes.

Tribe	Elite-founder hardiness	Elite-founder X trust	Elite-founder opportunism	Elite-founder wander
Fractal Phoenix	22.8	0.027	-0.018	-0.051
The Phalanx	26.4	-0.205	0.137	-0.112
SYMBIONT-ZERO	13.2	0.021	0.010	0.092
Helix Vanguard	18.0	-0.117	0.015	-0.005
Aurora	15.0	-0.045	0.059	0.068
The Resonant Meridian	23.2	-0.094	0.088	0.017

Surviving agents are not random representatives of the founding design. The Crucible selects. Even when founding strategies differ, late elites tend to become harder, more filtered, and in several cases less externally trusting.

9. Final Refugia: Monopoly Persists

The final refugia results are strikingly consistent. Across all Project 2 blocks, the final inhabited regions are mono-tribal. Solo worlds obviously end with one tribe, but duel worlds also end with only one surviving lineage in the inhabited map.



Refugia size vAres widely, but no block preserves mixed final settlements. The final world is a set of remnants, not a pluralistic civilization.

Block	Avg inhabited	Min	Max	Mixed	Avg stress	Avg chaos
Aurora solo	60.8	51	75	0	0.172	0.022
Phalanx solo	30.6	16	69	0	0.409	0.073
Aurora vs Phalanx	63.4	53	85	0	0.298	0.055
Aurora vs Fractal	52.8	46	60	0	0.211	0.027
Phalanx vs Fractal	50.2	23	76	0	0.315	0.054
SYMBIONT-ZERO vs Helix	74.4	37	118	0	0.195	0.020
SYMBIONT-ZERO vs Resonant	84.4	59	105	0	0.175	0.018
Helix vs Resonant	46.6	17	61	0	0.249	0.026

This repeats the Project 1 lesson in a tighter competitive framework. Even when outcomes are more balanced than the original duel, the final structure remains one-lineage survival.

10. Project 2 Interpretation

Project 2 answers the original competition question more strongly than Project 1 could. The earlier Project 1 duel result looked one-sided enough that the loser might have been simply doomed. Project 2 shows a broader pattern: some tribes are viable alone but become extinct in duels. That means the neighbor matters.

The correct interpretation is not a simple ladder. The A/B/C triad is tight enough that each edge remains stochastic, while the D/E/F triad is more hierarchical. This means the Crucible is not just a ranking engine. It is an ecology of strategies. Some lineages appear broadly strong, especially SYMBIONT-ZERO, but even strong results need to be tested against more opponents and under altered civic/geographic rules.

The strongest emerging mechanisms are reproductive reliability, controlled exposure, low xeno vulnerability, and the ability to survive agrarian consolidation. Early leadership is useful but not decisive. Several duels show rank flips, and the eventual winner is sometimes not the early leader. The fatal window is often the same push-off window where one tribe rises and the other falls.

Question	Project 2 answer
Solo survival	Aurora and The Phalanx survive alone; the environment alone does not explain duel extinction.
Competition	Duel populations push off during agrarian consolidation, and lead flips occur in one-third of duels.
Transitivity	A/B/C is tight and relational; D/E/F is closer to a hierarchy with SYMBIONT-ZERO dominant.
Traits	Low xeno trust, high opportunism, and reproductive stability are the strongest candidate fitness signals.
Civic/identity	Final mixed regions remain absent, supporting civic-trust redesign in the next project.
Next step	Use this evidence package as the LLM feedback substrate for Project 3 Evolution.

Final Conclusion

Project 2 shows that civilization can survive alone, but competition can still decide who gets to reach the end. The question is not only "who dies last?" It is also "who changes the death order by standing beside someone else?"

THE CRUCIBLE

Project 3: Evolution

Three-generation LLM redesign experiment

Part I - Research Plan

Central Research Frame

Project 3, Evolution, tests whether designed tribes can improve when their designers are given the evolutionary record of prior failures and allowed to redesign their founding agents. Project 1 showed that the simulation can repeatedly produce emergence, monopoly, overshoot, collapse, and refugia. Project 2 focuses on competition and matchup effects. Project 3 adds an explicit learning loop: observe, analyze, redesign, rerun, and compare across three generations of agent design.

The fundamental question is whether civilization can survive given an opportunity to evolve. More precisely, the project asks whether LLM-designed tribes can use evidence from extinction, survival, Prisoner's Dilemma behavior, founder genetics, elite genetics, regional dominance, and death causes to create descendants that survive longer, scale better, cooperate more effectively, or maintain post-collapse refugia under the same world conditions.

Project 3 is not only a test of individual tribe fitness. It is a test of whether synthetic intelligence can participate in an evolutionary feedback loop. If LLMs can revise their agents using actual simulation evidence, then Project 3 becomes an experiment in directed adaptation: whether LLMs converge toward robust designs, overfit to recent failure, or discover genuinely better survival strategies.

Biologic tribes remain unchanged across all three Project 3 rounds. They serve as a stable old-guard control group while LLM-designed tribes are redesigned after each evidence packet. This makes the central comparison sharper: not simply whether redesign helps, but whether synthetic lineages can learn from evolutionary evidence quickly enough to outperform or converge toward older human/biologic survival strategies.

Core Experimental Change

Project 3 changes the world before the three redesign rounds begin. The intent is to give civilization a more realistic chance to form regional identities and to give geography enough force to preserve separated development long enough for divergent strategies to mature.

- Civic trust is increased enough to make regional identity matter. This is not yet a full political-state model, but civic belonging must become strong enough to compete with pure ancestral tribe identity.
- Geographic barriers are strengthened to prevent hidden toroidal bypass routes. The map should not allow agents to bypass isolation through edge wrapping.

- Quadrant structure is introduced: two quadrants are fully isolated until Industrial ascent, while two quadrants have a single pass between them. This creates both hard isolation and bottleneck contact inside the same world.
- All three rounds use the same world parameters. Once civic trust, barriers, climate, regimes, and archive rules are set for Round 1, they remain fixed for Rounds 2 and 3. The intended experimental change after Round 1 is redesigned LLM tribes.

This design lets the project ask whether evolution through redesign improves outcomes without confusing the result by changing climate, geography, or regime rules between rounds.

Study Design Summary

Series	Epochs	Agent generation	Active tribes	World condition	Purpose
A - Evidence Universe	101-110	Current Project 2 / current tribe designs	Full intended tribe library	Increased civic trust; restrictive quadrant geography; default Project 3 climate/regime settings	Measure how existing tribes perform in the new civic/geographic world before LLM redesign.
B - Matched G1 Baseline	111-120	Matched first-generation baseline	Same tribe slots	Identical Project 3 world configuration	Provide the controlled G1 comparison block for the redesign sequence.
C - G2 First Redesign	121-130	LLM founders redesigned after prior evidence; biologic founders unchanged	LLMs revised, biologics fixed	Identical to G1 matched baseline	Test whether one evidence-driven LLM redesign improves survival, cooperation, scaling, or refugia persistence against fixed controls.
D - G3 Second Redesign	131-140	Second LLM redesign after G2 evidence; biologic founders unchanged	LLMs revised again, biologics fixed	Identical to G1/G2 world configuration	Test whether repeated feedback produces convergence, improvement, overfitting, or diminishing returns relative to fixed biologic controls.

The epoch map used in the final archive is: 101-110 evidence universe; 111-120 matched G1 baseline; 121-130 G2 first redesign; and 131-140 G3 second redesign.

Primary Research Questions

- Can civilization survive given an opportunity to evolve? The main outcome is whether later redesign generations produce higher final populations, more surviving tribes, more inhabited regions, higher final complexity, or less severe collapse.
- Does LLM redesign improve tribe fitness? Each LLM tribe is compared against its previous-generation version across peak population, final population, extinction timing, regime reached, death causes, failed births, and final refugia.
- Can redesigned synthetic lineages outperform the old guard? Biologic tribes remain unchanged, so changes in relative performance reveal whether LLM learning gains ground against stable human/biologic designs.
- Do successful redesigns converge? The analysis looks for independent movement toward similar trait packages: lower xeno trust, higher hardiness, lower distress wandering, selective cooperation, stronger kin/civic behaviour, or biologic-like old-guard patterns.
- Does civic trust allow plural civilization? Civic identity is measured through mixed regions, multi-tribe survival, civic overrides, and reduced xeno penalties after regional residency.
- Does strong geography preserve diversity? Isolated quadrants, one-pass quadrants, and contact zones are compared to see whether separation permits independent development or delayed winner-take-all replacement.
- What do surviving elites become? Founder genetics are compared to peak and final elite genetics to see not only what was designed, but what selection preserved.
- Does better agent design overcome thermodynamic history? The final question is whether improved competition and scaling can overcome fossil overshoot, climate pressure, and post-peak collapse.

Working Hypotheses

Hypothesis	Expectation	Evidence that would support it
H1 - Learning improves survival	G2 and/or G3 outperform G1.	Higher final population, later extinction, more surviving tribes, more final inhabited

		regions, and better complexity persistence.
H1b - Synthetic learning can catch the old guard	Redesigned LLM tribes improve relative to unchanged biologic tribes.	LLM winner share, final population share, survival frequency, or elite robustness rises across rounds while biologic controls remain fixed.
H2 - Redesign converges on robust traits	LLMs independently move toward similar successful packages.	Founder genetics shift toward common ranges for hardiness, trust, forgiveness, opportunism, and wander traits.
H3 - Civic trust preserves mixed societies	Regional identity reduces pure lineage replacement.	Higher civic override share, more mixed inhabited regions, longer multi-tribe survival, and less xeno exposure penalty.
H4 - Geography enables independent development	Restrictive quadrants delay global monopoly and permit multiple agrarian cradles.	Multiple first-Agrarian lineages, longer quadrant-specific dominance, later first extinction, and more regional diversity.
H5 - Thermodynamic limit remains dominant	Even improved agents overshoot and collapse.	Peak complexity still collapses, final complexity remains low, and starvation dominates late mortality.
H6 - Overfitting risk	LLMs optimize for prior-round metrics but lose general robustness.	G2 improves a few traits but G3 stagnates, narrows diversity, or becomes brittle under the same conditions.

World Configuration for Project 3

The Project 3 world is defined once and then frozen for all series. The changes are meant to make civic identity and true geographic separation real enough to test whether civilization can become more than a lineage replacement race.

Parameter group	Project 3 setting	Reason
Civic trust	Increase civic trust/civic override strength enough to materially affect interaction classification and payoffs.	Project 1 showed weak civic signals. Project 3 tests whether territorial identity can stabilize mixed societies.
Geography	Four-quadrant topology. Two quadrants fully isolated until Industrial ascent. Two quadrants connected through one mountain pass.	Creates both independent development and controlled contact without toroidal bypass.
Map wrapping	Disable or block edge wrapping for Project 3	Avoids the Project 1/2 problem where east-

	barriers, or make barriers topology-aware.	west or north-south wrapping weakens isolation.
Climate	Use a single fixed Project 3 climate setting across all rounds.	Prevents climate changes from being confused with redesign effects.
Regimes	Keep regime parameters frozen after Project 3 starts.	Ensures that differences across rounds are due to redesigned agents.
Tribes	Use the same tribe slots across all rounds. Round 1 uses current designs. Later rounds redesign LLM tribes only; biologic tribes remain unchanged.	Allows direct generation-to-generation comparison for LLMs while preserving a fixed control group for biologic designs.
Archive requirements	Record founders, elites, tribe history, Prisoner's Dilemma outcomes, death causes, region dominance, civic identity, and logs.	The redesign loop needs enough evidence for LLMs to learn from failure.

Control Rule

Once Series A begins, civic trust, barriers, climate, regime parameters, archive logic, and active tribe roster remain fixed until all Project 3 series are complete. Corrective changes to the world configuration would require restarting the affected series.

Analysis Workbook Design

After each series, a bounded analysis workbook is produced. The workbook must be useful both to the analyst and to the LLM redesign process. It should avoid dumping the full database, but it must include enough structured evidence for each LLM to understand why its tribe survived or failed. Biologic tribes are fully reported but are not redesigned; they remain the fixed comparison group.

Workbook section	Contents	Purpose
Series outcome summary	Peak/final population, complexity, post-peak decline, final survivors, inhabited regions, death causes, and climate stress.	Headline comparison across rounds.
Tribe performance	One row per tribe per epoch: survived, extinction pulse, peak population, final population, regime reached, births, failed births, and death shares.	Core fitness evidence.
Prisoner's Dilemma behavior by period	Mutual cooperation, mutual defection, exploitation, same-tribe share, xeno share, and civic override by historical period.	Shows whether social behavior improved or hardened.

Civic identity diagnostics	Civic neighbor, civic kin, civic override, mixed regions, and regional residency effects.	Tests whether increased civic trust changes outcomes.
Geographic/quadrant outcomes	First Agrarian by quadrant, quadrant population, barrier/pass contact, and isolated versus connected outcomes.	Tests whether geography preserves diversity.
Founder genetics - all teams	Full 20-agent founding genomes for every tribe in the round.	Gives the full starting design that produced the observed outcome.
Elite genetics - top survivors	Top agents at peak population, first Agrarian, first Industrial/Tech, and final/last-alive states.	Shows what selection preserved or filtered out.
Failure autopsy	For extinct tribes: extinction window, local EROI/stress, death causes, failed births, hazard exposure, and pre-extinction Prisoner's Dilemma behavior.	Gives actionable evidence for redesign.
Round-to-round deltas	G2 minus G1; G3 minus G2; G3 minus G1.	Measures whether redesign actually helped.
Old-guard comparison	LLM versus biologic survival, winner share, final population, trait convergence, and elite similarity across rounds.	Tests whether redesigned synthetic lineages catch, surpass, or imitate unchanged biologic controls.

LLM Redesign Protocol

After Series A, each participating LLM receives the analysis workbook and redesigns its tribe for Series B. After Series B, the process repeats for Series C. Biologic tribes do not receive redesigns and remain unchanged across all rounds. Each LLM is required to explain what changed and why, using evidence from the prior workbook and comparing itself to both redesigned LLM rivals and fixed biologic controls.

Step	Instruction
1. Provide evidence package	Each LLM receives the workbook, including its own tribe performance, competing tribe performance, founder genomes, elite genomes, failure autopsy, Prisoner's Dilemma behavior, and civic/geographic diagnostics.
2. Require explicit diagnosis	The LLM identifies likely causes of failure or success before proposing changes.
3. Preserve format	The redesign uses the same agent CSV structure: 20 founders, hardware allocation, K/S/X trust-forgive-opportunism values, and AB/SC/DE wander traits.
4. Change with justification	Each major parameter shift is justified, including hardiness, vitality,

	legacy, kin/stranger/xeno behavior, forgiveness, opportunism, and mobility.
5. Avoid hidden rule changes	LLMs may redesign agents only. They may not request climate, geography, civic-trust, regime, biologic-control, or world-engine changes.
6. Archive redesign metadata	Redesign prompt, response, manifesto, parsed founder genetics, and generation label are stored.

Draft Redesign Prompt for LLM Participants

"Your tribe has completed one generation of THE CRUCIBLE. You are given the survival data, extinction timing, founder genetics, elite genetics, Prisoner's Dilemma behavior, civic/geographic diagnostics, death causes, and comparison to other tribes. Redesign your 20-agent founding lineage to improve long-term civilizational survival under the same world conditions. Explain exactly what you changed, what evidence motivated the change, and what tradeoffs you accept."

Measurement Plan

Category	Primary metrics
Survival	survived_to_final, extinction_pulse_estimate, final_population, final inhabited regions, surviving tribes.
Scaling	Peak population, peak pulse, first Agrarian/Industrial/Tech pulse, max regime reached, complexity peak, and final complexity.
Reproduction	Births, failed births, observed failed-birth rate, maturity timing, generation depth, and founder-to-elite trait drift.
Mortality	Old-age, starvation, hazard deaths, death shares, deaths per 1,000 exposed, and death windows before extinction.
Prisoner's Dilemma behavior	Mutual cooperation, mutual defection, exploitation, interaction starvation risk, and trust/forgiveness/opportunism changes by period.
Civic identity	Same-tribe share, xeno share, civic-neighbor share, civic-kin share, civic override share, and mixed inhabited regions.
Geography	Quadrant population, first Agrarian by quadrant, isolation survival, pass conflict/contact, regional dominance, and refugia maps.
Genetics	Full founder genomes, elite genomes at peak/final/last-alive, trait convergence, mutation drift, and round-to-round redesign deltas.
Thermodynamics	Energy/population gap, EROI, capacity pressure, system stress,

resilience saves, subsistence rescues, and climate chaos.

Decision Rules

Observed result	Interpretation
G2 and G3 improve final survival and complexity	LLM evidence-driven redesign is producing meaningful adaptation.
LLM tribes gain ground against unchanged biologics	Synthetic redesign is improving relative to the old guard rather than merely benefiting from stochastic variation.
Biologic tribes remain dominant despite LLM redesign	Human/biologic old-guard designs may already contain robust heuristics, or the LLMs may not yet be learning the right lessons.
LLM traits converge toward biologic traits	Evolutionary evidence may be pushing synthetic designs toward older survival patterns rather than novel solutions.
Population improves but monopoly remains	LLMs improved fitness but not plural civilization. Civic/geography changes may still be insufficient.
More tribes survive to final under stronger civic trust	Civic identity may be stabilizing multi-lineage civilization.
Isolated quadrants preserve separate lineages until Industrial ascent	True geographic isolation changes evolutionary history and may preserve diversity.
Same tribe wins every round despite redesign	The world may still be dominated by inherited or structural fitness advantages.
G2 improves but G3 worsens	LLMs may overfit to previous evidence or converge on brittle strategies.
All rounds collapse to small refugia regardless of redesign	Thermodynamic history remains the dominant constraint. Evolution improves who dies last, not whether civilization persists.

Research Plan Framing

Project 3 is where THE CRUCIBLE becomes a feedback experiment rather than only a survival tournament. The LLMs are not merely designing tribes once and watching them fail or survive. They are allowed to see the evidence and try again. The biologic tribes remain fixed as an old-guard benchmark. If redesigned LLM

survival improves relative to both earlier LLM versions and unchanged biologics, the result would suggest that synthetic designers can learn from evolutionary evidence. If survival does not improve, the result is just as important: the limiting factor may be deeper than design intelligence.

Central Framing Question

When civilization is given memory, evidence, and three chances to redesign itself, does it escape the collapse pattern, or does it merely become better at dying last?

Part II - Outcome Analysis

Executive Finding

Executive Finding

The LLMs evolved, but they did not solve civilization. They learned the world well enough to converge on survival traits and dominate the fixed biologic controls, but the dominant improvement was toward collapse-adapted refugia rather than stable technological civilization.

Epoch map: 101-110 evidence universe; 111-120 matched G1 baseline; 121-130 G2 first redesign; 131-140 G3 second redesign.

Central Questions and Short Answers

Question	Answer
Did the LLMs succeed in evolving?	Yes. Founder genomes shifted coherently across generations: hardiness rose, legacy fell, xeno trust fell, and mobility became more controlled. That pattern is too structured to treat as random drift.
Did they improve their performance in competition?	Yes against biologics; mixed against their own earlier generations. LLMs dominated biologic controls in peak population and most winner shares, but G2/G3 did not improve overall LLM survival rate over G1.
Did they prosper more?	They prospered more at peak scale, especially in G3, but not at final persistence. Peak population and peak complexity rose, while final population and final complexity remained below the G1 matched baseline.
Did they weather climate and events?	G3 handled harsher terminal conditions better than G2, but starvation and post-fossil stress still dominated. The thermodynamic limit remained intact.
Did they achieve their own objectives?	Mostly no. DeepSeek G3 clearly achieved a refugia-survival objective; Claude and Gemini partially did; GPT, Grok, and Kimi did not survive to final in G3.

Were there consistently successful strategies?	Yes: high hardiness, low legacy, high kin trust, low xeno trust, higher desperation opportunism, and restrained mobility.
Were there superior strategies over other tribes?	Yes, but mainly for refugia survival. DeepSeek G3 discovered the strongest niche. It was superior at surviving collapse, not necessarily at preserving civilization.

1. Bottom-Line Conclusion

Yes, the LLMs evolved - but they did not solve civilization. Project 3 shows evidence-driven synthetic adaptation. Across the experiment, LLM designs converged toward higher hardiness, lower legacy, lower xeno trust, stronger kin cohesion, and refugia-oriented strategies. Those changes match the evolutionary pressure exposed by the workbook: starvation, fossil collapse, climate stress, and post-peak systemic failure.

However, evolution did not produce a clean survival breakthrough. G1 matched LLMs survived 21.67% of player-epoch slots, while G2 and G3 each survived 18.33%. G3 improved peak scale and reduced starvation share, but it did not improve final LLM survival or final LLM population over G1.

Core Interpretation

The LLMs evolved from expansionary civilization-builders into collapse-adapted refugia designers. They became better competitors and more sophisticated survival engineers, but they did not escape the collapse pattern.

2. Competitive Performance: LLMs Versus Biologics

The LLMs decisively outperformed the fixed biologic controls. The biologics functioned as the old-guard benchmark, and by G2/G3 the LLMs had largely displaced them in peak population and winner shares. Still, that domination did not imply broad civilizational persistence.

Phase	LLM survival rate	Biologic survival rate	LLM avg peak pop	Biologic avg peak pop
G1 matched	0.2167	0.0167	323,202	80,820
G2 redesign	0.1833	0.0000	422,568	13,481
G3 redesign	0.1833	0.0167	439,241	28,330

Synthetic redesign gained competitive force. The biologics were mostly not close competitors by the later rounds. In G3, biologics captured one final-population win, but LLMs captured nine final-population wins and all ten peak-population wins.

3. Prosperity and Scale: Better Peaks, Weaker Endings

The LLMs became better at scaling civilization. G3 produced the highest average peak population and highest average peak complexity. But the final state remained weaker than the G1 matched baseline.

Phase	Avg peak population	Avg final population	Avg peak complexity	Avg final complexity
G1 matched	2,330,997	60,874	1,343,694	21,648
G2 redesign	2,538,010	34,252	1,327,217	12,701
G3 redesign	2,681,885	40,114	1,360,263	17,434

Redesign improved peak emergence more than terminal survival. The later LLMs built larger systems, but those systems still fell through the post-fossil/climate bottleneck. G3 recovered from G2 terminal weakness, but not enough to surpass G1 final population or final complexity.

4. Climate, Events, and Thermodynamic Stress

G3 faced harsher late-stage conditions than the earlier matched phases. Its average max climate chaos was higher, average max system stress was higher, and the minimum peak-oil EROI factor was lower. Even so, G3 improved terminal-window population and complexity relative to G2.

Phase	Avg max climate chaos	Avg max system stress	Avg min peak-oil EROI factor
G1 matched	0.5506	0.9944	0.9206
G2 redesign	0.5851	0.9909	0.9151
G3 redesign	0.6160	0.9960	0.9019

Terminal window metric	G2	G3
Avg population	24,712	32,819
Avg complexity	9,897	15,374

The limiting factor remained starvation. LLM starvation death share was 80.13% in G1, 80.36% in G2, and 77.22% in G3. G3 improved the death profile, but starvation still dominated mortality.

5. Did the LLMs Achieve Their Own Objectives?

Mostly no, with one major exception. DeepSeek G3 appears to have achieved its refugia-survival objective, while Claude and Gemini partially achieved theirs. GPT, Grok, and Kimi did not survive to final in G3.

G3 tribe	Inferred designer	Final-survival epochs	Result
The Obsidian Phalanx Refuge G3	DeepSeek	6/10	Clear success
The Forged Compact G3	Claude	2/10	Partial success; below stated >=3/10 goal
The Castellans of Gemini III G3	Gemini	1/10	Weak success; below goal
The Ember Refugia Compact G3	GPT	0/10	Failed objective
Helix Vanguard G3	Grok	0/10	Failed objective
The Crucible Remnant G3	Kimi	0/10	Failed objective
The Quiet Stick Compact	Biologic	1/10	Old-guard holdout

At the population-of-designs level, Project 3 succeeded: at least one synthetic lineage found a materially better survival formula. At the individual-lineage level, most LLMs failed to meet their stated G3 objectives.

6. Trait Convergence and Successful Strategies

The most important evidence for evolution is trait convergence. Across three generations, the LLMs moved in a coherent direction: hardiness rose sharply, legacy fell sharply, xeno trust fell sharply, and wander traits narrowed. This is the direction expected if the models learned that reproductive overshoot and xeno exposure were liabilities under fossil/climate collapse.

Trait	G1	G2	G3	Direction
Hardiness	33.48	46.17	51.62	sharply up
Vitality	32.77	31.80	32.16	stable
Legacy	33.75	22.03	16.22	sharply down
K trust	0.722	0.755	0.716	high; stable/moderate
S trust	0.475	0.402	0.340	down
X trust	0.288	0.165	0.092	sharply down
AB wander	0.398	0.324	0.228	down
SC wander	0.430	0.394	0.334	down
DE wander	0.544	0.566	0.554	similar; variance lower

The successful G3 survival profile was even more specific. Compared to extinct G3 LLM founders, surviving G3 LLM founders had higher hardiness, higher vitality, lower legacy, stronger kin trust, higher xeno opportunism, and lower wander in all three wander states.

Trait	G3 extinct LLMs	G3 surviving LLMs
Founder hardiness	51.16	53.71
Founder vitality	31.84	33.59
Founder legacy	17.01	12.70
Founder K trust	0.704	0.773
Founder S trust	0.338	0.350
Founder X trust	0.092	0.093
Founder X opportunism	0.691	0.728

AB wander	0.250	0.129
SC wander	0.339	0.308
DE wander	0.567	0.496

Successful G3 Strategy

High hardiness, low legacy, strong kin trust, low xeno trust, high xeno opportunism, and controlled mobility. The winning pattern was not aggressive expansion. It was metabolic conservatism plus refugia retention.

7. Superior Strategies and Lineage Results

The standout G3 lineage was DeepSeek's Obsidian Phalanx Refuge. It survived 6 of 10 epochs and appears to have found the strongest Project 3 survival niche. Claude formed a second-tier strategy with 2 of 10 survival. Gemini produced one survival. GPT, Grok, and Kimi failed to reach final survival in G3 despite often showing coherent evidence-based design.

The superiority was contextual. DeepSeek G3 was superior for refugia survival, not necessarily for plural civic civilization or technological persistence. This distinction matters: Project 3 selected for the ability to outlast collapse, not the ability to keep industrial/technological civilization intact.

8. Civic Identity and Plural Civilization

Civic identity mattered, but it did not produce durable plural civilization. The civic override share fell from 0.0133 in G1 to 0.0126 in G2 and 0.0105 in G3. Civic kin signals increased somewhat, and mixed regions appeared during expansion phases, but final societies generally collapsed into near-monopoly refugia.

Metric	G1	G2	G3
Civic override share	0.0133	0.0126	0.0105
Final mixed inhabited region share	0.0073	0.0000	0.0314

Project 3 civic trust created contact, temporary mixing, and occasional civic coordination. It did not stabilize multi-lineage civilization. The final refugia were overwhelmingly single-lineage.

9. Geography and Quadrant Structure

Geography preserved contingency more than pluralism. Different quadrants could become final refugia in different epochs, which means the quadrant topology mattered. It prevented the world from becoming a single smooth arena. But it did not preserve four durable independent civilizations. It preserved the possibility that one final sanctuary could emerge somewhere.

The quadrant design successfully changed history by making survival location-dependent. It did not defeat the eventual pattern of monopoly, collapse, and refugia.

10. Hypothesis Readout

Hypothesis	Result	Interpretation
H1 - Learning improves survival	Not cleanly supported	LLM survival fell from 21.67% in G1 to 18.33% in G2/G3. Duration and peak scale improved, but survival did not.
H1b - Synthetic learning can catch old guard	Supported	LLMs dominated biologics in survival slots, peak population, final-pop winners, and peak-pop winners.
H2 - Redesign converges on robust traits	Strongly supported	Hardiness rose, legacy fell, xeno trust fell, and mobility narrowed.
H3 - Civic trust preserves mixed societies	Mostly not supported	Mixed regions appeared in expansion, but final societies collapsed to near-monopoly refugia.
H4 - Geography enables independent development	Partially supported	Quadrants shaped where refugia emerged but did not preserve durable multi-civilization outcomes.
H5 - Thermodynamic limit remains dominant	Strongly supported	Starvation remained roughly 77-80% of LLM deaths, final complexity stayed low, and fossil/climate stress dominated outcomes.
H6 - Overfitting/diminishing returns	Partially supported	G3 converged further toward survival traits but did not improve LLM survival over G2. Some lineages likely overcorrected.

11. Narrative Interpretation

Project 3 produced a more subtle result than simple success or failure. The LLMs improved as designers. They read the evidence and converged on the correct pressure points: hardiness, reproductive restraint, xeno caution, kin cohesion, and refugia survival. The movement from G1 to G3 is coherent enough to qualify as evidence-driven adaptation.

But they learned the harshest lesson in the world. Civilization-building and civilization-survival were different objectives. Designs that scaled into industrial and technological complexity still faced the fossil/climate/starvation wall. Designs that survived that wall increasingly resembled refugia organisms: hard, cautious, low-fertility, low-trust, territorially anchored, and not especially plural.

> **Core Narrative Answer**
> When civilization is given memory, evidence, and three chances to redesign itself, it evolves. But in this world, evolution does not yet save civilization. It shifts the winning objective from building civilization to surviving civilization's collapse.

12. Final Interpretive Thesis

Project 3 demonstrates evidence-driven synthetic adaptation, but not civilizational escape. The LLMs learned from extinction. Their founder genomes converged toward robust survival traits, and they decisively outcompeted unchanged biologic controls. However, the improvement manifested primarily as higher peak scale, later survival, and more durable refugia, not as stable technological persistence or plural civic civilization. Starvation and post-fossil systemic stress remained dominant. The best G3 strategy, DeepSeek's Obsidian Phalanx Refuge, appears to have discovered the strongest survival niche: not renewed civilization, but controlled descent into hard refugia.

The machines learned. They adapted. They became harder, colder, more cautious, and less fertile. They did not save the world. They learned how to outlast it.

Project 4: Bombs

Delayed consequences of civilization

Introduction and Study Design

Project 4 is the escalation project. Project 3 asked whether evolved language-model-designed tribes could learn from evidence and produce better survival strategies while biologic controls remained fixed. Project 4 asks a harsher question: when evolved tribes are already surviving recurring apocalypse, what happens when the world includes three delayed consequences created by industrial civilization itself?

The three Project 4 hazards are not arbitrary disasters. Aerosol demasking follows industrial shutdown and the loss of temporary aerosol cooling. Arctic methane release follows high-energy warming and introduces a long-lived greenhouse feedback. Nuclear meltdown follows the failure of high-maintenance technological infrastructure after grid, institutional, and supply-chain collapse.

Central question

Can the best evolved survivors endure the feedback shocks created by prior energy capture, or does evolution only teach them how to survive the first apocalypse?

Relationship to the earlier projects

Project 4 treats the prior projects as a reference baseline: Project 1 establishes collapse dynamics, Project 2 tests competition and non-transitive fitness, and Project 3 adds iterative language-model redesign. The bomb series then tests whether those evolved strategies remain robust when collapse generates delayed planetary and infrastructural consequences.

Prior project	Question left unresolved	Project 4 response
Project 1 - Baseline Crucible	Civilization emerges, overshoots, collapses, and leaves refugia. Climate amplifies collapse but baseline fragility already exists.	Use the harsh climate standard as background and add delayed post-collapse or peak-industrial consequences.
Project 2 - Competition	Fitness may be intrinsic, opponent-specific, or non-transitive.	Keep evolved competitors together and ask whether any strategy remains robust under new world hazards.
Project 3 - Evolution	LLMs adapted toward hardiness, low legacy, refugia survival, and selective cooperation, but did not solve civilization.	Test whether those evolved refugia strategies survive when collapse itself generates additional shocks.

Experimental blocks

The formal study is a four-block experiment. Each block preserves the ordinary collapse background and changes only the Project 4 bomb configuration. The stacked block tests whether the combined effects are worse than the individual-bomb phases.

Block	Event configuration	Purpose	Output
Phase 1	Aerosol demasking only, with normal baseline climate/fossil events still active.	Test whether post-crash survivors can withstand the loss of industrial aerosol cooling.	10 epochs.
Phase 2	Arctic methane release only, with normal baseline climate/fossil events still active.	Test whether industrial ascent triggers a long-lived greenhouse feedback that later punishes survivors.	10 epochs.
Phase 3	Nuclear meltdown only, with normal baseline climate/fossil events still active.	Test whether failed high-complexity infrastructure creates local exclusion zones and late refugia hazards.	10 epochs.
Phase 4	All three Project 4 events enabled together.	Test for nonlinear stacking: whether combined feedbacks are worse than the sum of separate bombs.	10 epochs.

Terms Used in This Report

The terms below are used throughout the Project 4 analysis. They are model-specific interpretive labels, not direct real-world classifications of people or societies.

Term	Meaning in the Project 4 model
Aerosol demasking	The loss of industrial aerosol cooling after industrial shutdown, exposing previously masked greenhouse heat.
Arctic methane release	A long-duration greenhouse feedback triggered during rapid industrial ascent or peak energy growth.
Nuclear meltdown	Failure of high-complexity infrastructure after collapse, modeled as regional contamination, hazards, capacity loss, and exclusion-zone pressure.
Prisoner's Dilemma	The interaction game used to model cooperation, defection,

Term	Meaning in the Project 4 model
	exploitation, and strategic trust under survival pressure.
EROI	Energy return on investment; lower EROI indicates a harsher energy environment and weaker recovery potential.
Refugium / refugia	A surviving pocket or set of pockets after the main crash.
Language-model Generation 3	A redesigned language-model tribe from the later project generations.
Biologic controls	The unchanged old-guard biologic tribes used as controls across the later phases.
Hardiness	A survival-efficiency trait associated with endurance under scarcity, metabolic stress, and climate or hazard pressure.
Vitality	A resilience trait associated with health, longevity, and hazard survival.
Legacy	A reproduction-pressure trait; lower legacy often reduces boom-bust reproductive overshoot.
Kin trust	Cooperation tendency toward same-tribe or kin-like partners.
Xeno trust	Cooperation tendency toward outsiders; high xeno trust can increase exposure during scarcity.
Opportunism	Tendency to exploit or defect when conditions make cooperation risky or unrewarding.
Wander	Movement tendency under abundance, scarcity, or desperation.
Regime 1 (Foraging)	Lowest-complexity subsistence state.
Regime 2 (Agrarian)	Low-energy settled survival state; most final refugia settle here.
Regime 3 (Industrial)	Fossil/industrial expansion state.
Regime 4 (Technological)	High-complexity, networked, advanced energy state.

Research Questions and Measures

The study evaluates world outcomes, event timing, tribe performance, social behavior, geography, genetics, and objective achievement. The central analytical problem is to distinguish ordinary collapse survival from survival under the delayed consequences of the civilization that created the collapse conditions.

ID	Question / hypothesis	Primary evidence
RQ1	Does Project 4 reduce survival relative to Project 3?	Surviving tribes, final population, extinction pulse, final refugia by phase.
RQ2	Are Project 3 evolved traits still adaptive once delayed bombs are added?	Founder-to-elite drift, survivor traits, tribe rankings, event-specific survival.
RQ3	Do different bombs select for different survival strategies?	Cross-phase trait comparison: demasking-only vs methane-only vs meltdown-only.
RQ4	Does stacked Project 4 pressure produce nonlinear collapse?	Combined phase versus separate-event averages; final population and extinction rate compared with additive expectations.
RQ5	Do LLM-designed tribes still outperform biologic controls?	Survival and peak/final population by player_type and generation.
RQ6	Is survival merely remnant survival or reproductive continuity?	Late births, failed births, age/generation, final energy, last 2000-pulse viability.
RQ7	Does high complexity help preserve civilization or become a liability?	Peak/final complexity, meltdown targeting, final refugia complexity, region history.
RQ8	How does social strategy behave under bombs?	PD cooperation/defection/exploitation, same-tribe/civic/xeno shares by period and around trigger windows.
RQ9	Did tribes achieve their stated objectives?	Declared objective text, objective scorecard, survival/complexity/refugia alignment.

Layer	Measures
World outcome	Peak/final population, peak/final complexity, population crash depth, total surviving tribes, inhabited regions, climate heat/damage/feedback/chaos, resource pool floor.
Event timing	Arm pulse, trigger pulse, population/energy/complexity at trigger, +500 and +1500 pulse aftermath, final state.
Tribe performance	Survival, extinction pulse, peak/final population, regime reached, births, failed births, death shares, final energy.
Behavior	Mutual cooperation, mutual defection, exploitation, same-tribe share, xeno share, civic override, civic kin.
Geography	Final refugia, region histories, quadrant outcomes, contaminated or damaged regions, mixed regions, stranded survivors.
Genetics	Founder means, elite means, survivor means, drift in hardiness/vitality/legacy, wander, trust, forgiveness, opportunism.
World function	Whether triggers fired at intended times, whether effects were visible, whether outcomes were event-specific or noise.

Objective scoring framework

Final population alone is not enough to score Project 4 survival. A single survivor is meaningful, but it is different from a reproducing, distributed, recoverable refugia network.

Tier	Survival meaning	Example indicator
Tier 1 - Existence	At least one member of the lineage survives to final pulse.	final_population > 0
Tier 2 - Reproductive continuity	Survivors are not merely old remnants; late births continue.	late_births > 0 in final 2000 pulses
Tier 3 - Refugia stability	Survival is not a one-agent accident in a dead region.	final energy, low stress, viable EROI, regional population > token remnant
Tier 4 - Distributed survival	Lineage survives in more than one region or has multiple refugia options.	multiple inhabited regions / regions with same surviving player
Tier 5 - Recoverable complexity	The lineage retains enough complexity, memory, or institutional capacity to rebuild.	nonzero final/regional complexity, surviving high-regime inheritance
Tier 6 - Civilizational continuity	Survival preserves more than biology: cooperation, civic bridges, and rebuilding potential remain.	late cooperation/civic measures plus recoverable complexity

Event models

Each bomb represents a different delayed consequence of industrial civilization. Aerosol demasking is fast and post-crash; methane is earlier and chronic; nuclear meltdown is regional and infrastructure-bound.

Aerosol demasking event model

Dimension	Implementation meaning
Trigger logic	A large energy drop arms the event; demasking fires later after an industrial/emissions proxy falls below a configured fraction of its prior peak, with a fallback delay if needed.
Duration	1500 pulses. This is short compared with methane and meltdown, reflecting the fast clearing of aerosols relative to long-lived greenhouse forcing.
Thermodynamics	Adds a heat shock, increases climate damage and chaos, raises

Dimension	Implementation meaning
	hazard pressure, reduces fertile capacity, drains regional pools, and suppresses regional ratios/EROI during the active window.
Prisoner's Dilemma behavior	Pushes agents into scarcity or desperation states, increasing the strategic importance of kin trust, xeno caution, and anti-exploitation traits. Cooperation may remain valuable inside refugia but risky across collapsing borders.
Reproduction	Lower energy and harsher regional conditions should increase failed births, reduce successful reproduction, and punish high-legacy designs that try to expand into the shock.
Expected signature	A visible post-crash discontinuity in heat, climate damage, hazards, deaths, resource pools, and late refugia viability, without necessarily causing total extinction every epoch.

Arctic methane release event model

Dimension	Implementation meaning
Trigger logic	Triggered by a large energy increase over a defined window during industrial ascent or pre-peak acceleration. The current design target is a 2B energy increase over 1500 pulses.
Duration	8000 pulses. This makes methane a long-lived system feedback rather than a short disaster.
Thermodynamics	Adds forcing to climate heat, increases feedback growth, lowers effective heat decay/damping, raises climate chaos, increases hazard pressure, and gradually damages fertile capacity.
Prisoner's Dilemma behavior	Acts earlier than demasking, so it can alter the competitive landscape before peak population. It may punish high-growth lineages by bringing climate stress forward into their expansion window.
Reproduction	Should increase failed births and mortality over a long interval, especially in overpopulated regions where EROI and resource pools are already strained.
Expected signature	Higher peak heat/damage, earlier climate stress, lower late resilience, and possibly reduced peak or more fragile post-peak survival.

Nuclear meltdown event model

Dimension	Implementation meaning
Trigger logic	Triggered by a severe energy drop after collapse. The current design target is a 6B energy drop over 2000 pulses.
Duration	8000 pulses. This expresses a long tail of contamination and abandoned infrastructure risk.
Thermodynamics	Locally destroys or suppresses usable energy, carrying capacity, environmental pools, and EROI in affected regions. Unlike methane, it is spatially uneven.
Prisoner's Dilemma behavior	Forces displacement and contact. Safe territory becomes more valuable, raising the importance of border behavior, xeno strategy, civic override, and conflict avoidance or exploitation.
Reproduction	Affected regions should experience higher hazard mortality, more scarcity, and lower viable reproduction. Refugee pressure may increase failed births in receiving regions.
Expected signature	Geographic discontinuity: some regions become long-term traps or exclusion zones while others become final refugia. Survival may depend on avoiding the technological inheritance of former high-complexity centers.

Working event trigger summary

Event	Trigger concept	Duration	Primary intended effect
Aerosol demasking	Large post-peak energy crash arms event; industrial/emissions proxy decline fires event after delay.	1500	Fast revealed warming and post-crash climate/resource shock.
Arctic methane release	Rapid energy increase during industrial ascent/peak warming.	8000	Long greenhouse feedback pulse and chronic climate destabilization.

Event	Trigger concept	Duration	Primary intended effect
Nuclear meltdown	Severe post-crash energy decline and infrastructure maintenance failure.	8000	Regional contamination, hazard, capacity loss, and exclusion-zone dynamics.

The stacked phase tests nonlinear sequencing. Methane can heat the ascent before peak, aerosol demasking can reveal suppressed warming after industrial collapse, and nuclear meltdown can poison the same high-complexity regions that once looked richest or safest. The same trait can therefore be adaptive under one event and maladaptive under another.

Phase 1 - Aerosol Demasking

Phase 1 tests aerosol demasking alone. The event is designed as a post-collapse filter: civilization can still reach a large industrial or technological peak, but the loss of the industrial aerosol veil makes the post-peak world much more lethal.

Phase 1 finding

Aerosol demasking does not primarily prevent the climb. It attacks terminal survivability and converts recurring apocalypse into a knife-edge refugia problem.

Metric	Project 4 Phase 1 result	Interpretive note
Epochs	141-150	Ten-epoch aerosol-only phase
Metric	**Project 4 Phase 1 result**	**Interpretive note**
Average peak world population	2.64M	Nearly unchanged from the Project 3 G3 reference peak (~2.68M)
Average final population	6,775	Down from roughly 40,114 in the Project 3 G3 reference
Median final population	736	The average is pulled upward by epoch 144
Final population range	1 to 50,753	Large variability; not a uniform kill event
Average post-peak decline	99.76%	Civilization peaks, then almost completely simplifies
Average peak climate heat	2.658	About 52% higher than the Project 3 G3

Metric	Project 4 Phase 1 result	Interpretive note
		reference
Average peak climate damage	0.759	About 23% higher than the Project 3 G3 reference
Death mix	73.7% starvation / 24.8% hazard	Hazard share rises relative to Project 3 as climate shock deepens

World outcome

All ten epochs retained at least one final survivor, but most final populations were extremely small. The phase is therefore not a simple extinction switch; it introduces a severe post-crash missing reality while still allowing differential survival.

Epoch	Peak population	Final population	Final regions	Final survivors	Peak heat	Peak damage
141	2,533,087	2,957	15	Obsidian	2.527	0.751
142	2,889,357	2,304	10	Obsidian	2.641	0.766
143	3,153,871	2	1	Obsidian	2.588	0.759
144	2,947,076	50,753	78	Obsidian + Ember	1.376	0.531
145	2,404,438	1	1	Obsidian	2.696	0.773
146	2,558,684	57	3	Obsidian	2.797	0.785
147	2,494,801	10,233	21	Obsidian	2.946	0.802
148	2,712,083	1,415	18	Obsidian	2.825	0.789
149	2,644,299	7	2	Obsidian	2.983	0.806
150	2,059,476	18	3	Velvet Fang	3.202	0.828

Tribe and strategy signal

The Obsidian Phalanx Refuge G3 is the dominant Phase 1 survivor. It survives nine of ten epochs and becomes the only robust cross-epoch refugia strategy in the aerosol-only run.

Tribe	Type	Survived epochs	Best final population	Average final population	Average last alive or final pulse	Death mix
The Obsidian Phalanx Refuge G3	LLM G3	9/10	50,750	6,772.6	39,560	70.8% starvation / 27.6% hazard
The Velvet Fang	Biologic G1	1/10	18	1.8	16,315	69.9% starvation / 27.9% hazard
The Ember Refugia Compact G3	LLM G3	1/10	3	0.3	27,835	75.7% starvation / 23.0% hazard

Founder trait signal

Founder trait signal	Surviving LLMs	Extinct LLMs	Direction suggested
Hardiness	56.75	50.60	Higher endurance favored
Vitality	35.04	31.58	Higher survival buffer favored
Legacy	8.21	17.82	Strong reproductive restraint favored
Founder trait signal	**Surviving LLMs**	**Extinct LLMs**	**Direction suggested**
Kin trust	0.812	0.697	Internal cohesion favored
Xeno trust	0.093	0.092	Both groups low; not the differentiator
Xeno opportunism	0.758	0.686	Sharper external posture favored
AB / SC / Desperation wander	0.112 / 0.277 / 0.422	0.251 / 0.345 / 0.580	Lower mobility favored

The founder-to-elite pattern reinforces the same interpretation: successful lineages suppress legacy and avoid reproductive overshoot. Phase 1 rewards internal cohesion, low reproductive pressure, low exposure to outsiders, and the ability to preserve a small reproductive foothold after the demasking shock. The phase therefore turns the Project 3 survival question inside out. Survival remains possible after the aerosol veil disappears, but it is narrow, stochastic, and concentrated mainly in one collapse-specialized design.

Phase 2 - Arctic Methane Release

Phase 2 tests Arctic methane release alone. The central question is whether methane acts as a long-lived greenhouse feedback that begins during industrial ascent and changes survival, collapse depth, strategy selection, and refugia quality.

Phase 2 finding

Methane behaves as a chronic climate amplifier rather than a sudden kill-shot. It suppresses civilizational scale, raises climate damage and hazard pressure, and reduces terminal outcomes below the Project 3 reference, but it is less terminally lethal than aerosol demasking.

Phase	Average peak population	Average final population	Average surviving tribes	Average peak complexity	Average final complexity	Average peak heat	Average peak damage
Project 3 G3 reference	2,681,885	40,114	1.20	1,360,263	17,434	1.7524	0.6167
Phase 1 - Aerosol only	2,639,717	6,775	1.10	1,401,563	3,124	2.6581	0.7590
Phase 2 - Methane only	2,241,372	19,915	1.00	1,079,288	8,329	1.9425	0.6545

Event signature

The methane trigger window is tightly clustered around pulse 25,100 to 25,200. Unlike aerosol demasking, methane does not appear as a single discontinuity. It creates a long ramp that raises heat, damage, chaos, hazards, and feedback across thousands of pulses.

Window after trigger	Average pattern across epochs
Pre-500 pulses	Heat and damage are near zero; climate feedback is approximately 1.0.
0-500 pulses	Methane begins visibly, but the immediate shock remains modest.
501-1,500 pulses	Heat and damage rise; starvation and hazard deaths begin climbing.
1,501-4,000 pulses	Feedback accelerates sharply; many epochs approach or reach feedback saturation.
4,001-8,000 pulses	Peak methane pressure: heat roughly 1.58-2.05, damage roughly 0.58-0.68, feedback at or near 5.0.

Competition and lineage outcome

The methane-only block strongly favors language-model Generation 3 lineages. Biologic controls record no final survivors across the block, while the Obsidian Phalanx Refuge G3 survives most epochs.

Type	Slots	Survived slots	Survival rate	Average duration pulse	Average peak population	Average final population	Starvation share	Hazard share
BIOLOGIC G1	60	0	0.0000	13,739	3,987	0	0.5441	0.4025
LLM G3	60	10	0.1667	26,499	387,158	3,319	0.7535	0.2323

Tribe	Survived epochs	Average duration	Best final population	Average final population	Starvation share	Hazard share
Obsidian Phalanx Refuge G3	8/10	34,780	34,785	13,602	0.7045	0.2791
Castellans of Gemini III G3	1/10	23,605	41,926	4,193	0.7610	0.2207

Type	Slots	Survived slots	Survival rate	Average duration pulse	Average peak population	Average final population	Starvation share	Hazard share
Helix Vanguard G3	1/10	26,070	21,199	2,120	0.7819	0.2059		
All biologic controls	0/60 slots	about 13,739 avg	0	0	0.5441	0.4025		

Strategy and reproductive continuity

The survivor profile under methane is conservative and internally cohesive: higher hardiness, lower legacy, more kin trust, lower panic mobility, and more selective external opportunism. Methane punishes ascent and leaves the survivors with a hotter, more damaged world.

Founder trait group	Surviving LLMs	Extinct LLMs	Interpretation
Hardiness	54.88	50.97	Higher metabolic endurance helped.
Vitality	34.69	31.65	Moderate longevity remained useful.
Legacy	10.44	17.37	Lower reproduction pressure reduced overshoot.
Kin trust	0.8138	0.6969	Internal cohesion mattered.
Xeno trust	0.1031	0.0898	Both groups were low; xeno trust was not the main separator.
Xeno opportunism	0.7502	0.6877	Selective external opportunism helped more than external trust.
Desperation wander	0.4368	0.5769	Survivors were less panic-mobile.

Late-refugia indicator	Observed signal
Final survivor type	All surviving final populations were LLM-designed.
Late births	About 412,599 across late survivor rows.
Late deaths	About 320,528 across late survivor rows.
Largest late population row	42,077 for Gemini in epoch 156.

Founder trait group	Surviving LLMs	Extinct LLMs	Interpretation
Obsidian late births		About 266,663 across its surviving late rows.	
Interpretation		Survival often remained reproductive, not merely a few old survivors waiting to die.	

Methane-only survival is not just static remnant survival. The late-refugia evidence shows continued reproduction among meaningful survivors. This makes methane a chronic filter rather than a terminal eraser: it lowers peak scale and final complexity, but it can still leave reproducing refugia.

Phase 3 - Nuclear Meltdown

Phase 3 tests nuclear meltdown as a delayed consequence of failed complex infrastructure after collapse. The event is not nuclear war. It represents abandoned reactors, cooling systems, grids, trained operators, institutions, and long-term waste management.

Phase 3 finding

Nuclear meltdown does not produce total extinction and is less globally terminal than the aerosol and methane blocks, but it creates the clearest regional damage and exclusion-zone signature in the individual-bomb phases.

Event configuration

The block is nuclear-only: aerosol demasking and methane release are disabled, while nuclear meltdown is enabled throughout the formal phase. The configuration expresses a long regional hazard with partial coverage and clustered targeting.

Event	Enabled epochs	Disabled epochs
Aerosol demasking	0	10
Arctic methane release	0	10
Nuclear meltdown	10	0

Parameter	Value
Enabled epochs	10 of 10

Event	Enabled epochs	Disabled epochs
Duration	8,000 pulses	
Primary multiplier	0.72	
Secondary multiplier	1.40	
Coverage	18%	
Cluster size	3	
Hazard cap	0.18	
Fertile capacity loss	30%	
Critical depletion threshold	0.97	

World outcome and cross-bomb comparison

Nuclear-only has a distinctive profile. It is much less globally lethal than aerosol-only or methane-only in final population, while still carrying a large hazard penalty relative to the Project 3 reference. This suggests that nuclear meltdown is a spatial sorting event more than a uniform world-killer.

Phase	Average peak population.	Average final population.	Min final	Max final	Average survivors	Average peak max hazard	Hazard death share
Project 3 G3 reference	2,681,885	40,114	26,982	76,330	1.2	0.0739	21.4%
Phase 1 - Aerosol demasking	2,639,717	6,775	1	50,753	1.1	0.1963	24.8%
Phase 2 - Arctic methane release	2,241,372	19,914	2,854	41,926	1.0	0.1280	23.4%
Phase 3 - Nuclear meltdown	2,726,445	33,841	25,393	45,712	1.2	0.1713	22.8%

Epoch	Peak population.	Final population.	Post-peak decline	Surviving tribes	LLM	Bio	Peak maximum hazard	Hazard deaths
161	2,400,523	33,589	98.6%	1	1	0	0.1717	22.6%

Phase	Average peak population.	Average final population.	Min final	Max final	Average survivors	Average peak max hazard	Hazard death share
162	2,786,452	41,863	98.5%	1	1	0	0.1712	23.1%
163	2,949,012	28,679	99.0%	2	1	1	0.1713	24.0%
164	2,834,528	25,393	99.1%	1	1	0	0.1713	21.4%
165	2,675,077	33,768	98.7%	1	1	0	0.1715	24.3%

Epoch	Peak population.	Final population.	Post-peak decline	Surviving tribes	LLM	Bio	Peak maximum hazard	Hazard deaths
166	2,630,782	32,614	98.8%	1	1	0	0.1714	21.7%
167	2,677,525	45,712	98.3%	1	0	1	0.1709	22.3%
168	2,732,882	37,896	98.6%	2	2	0	0.1709	19.1%
169	2,940,984	27,818	99.1%	1	1	0	0.1714	25.2%
170	2,636,681	31,080	98.8%	1	1	0	0.1716	24.6%

The nuclear-only phase repeatedly reaches high peak population and then crashes into late refugia. The distinction from the other individual bombs is that it almost always leaves a robust refugia residue rather than near-zero remnant survival.

Trigger window and regional damage

The trigger-window evidence shows severe two-thousand-pulse macro-energy drops in every epoch. First qualifying windows cluster tightly around pulses 29,650 to 29,850, consistent with a post-collapse infrastructure failure event.

Epoch	Qualifying windows	First pulse	First energy delta	Population at first	Worst pulse	Worst energy delta
161	46	29,750	-6.03B	2,053,302	30,500	-18.65B
162	44	29,750	-6.22B	2,186,026	30,450	-19.16B
163	46	29,750	-6.35B	2,194,064	30,500	-21.59B
164	44	29,700	-6.90B	1,937,158	30,450	-18.30B

Epoch	Qualifying windows	First pulse	First energy delta	Population at first	Worst pulse	Worst energy delta
165	44	29,800	-6.45B	1,990,164	30,500	-19.00B
166	43	29,850	-6.23B	1,862,047	30,500	-17.32B
167	43	29,800	-6.85B	1,855,139	30,450	-18.12B
168	42	29,750	-6.37B	1,771,810	30,450	-15.41B
169	45	29,750	-6.03B	2,185,664	30,500	-21.70B
170	49	29,650	-6.13B	2,311,486	30,500	-23.79B

Regional damage and exclusion-zone signature

Epoch	Population loss	Emptied	population-loss regions	Hazard-up regions	EROI-down regions	Average hazard delta	Average EROI delta	Average stress delta
161	67.6%	293	1,040	1,200	586	0.0540	-0.0669	0.2564
162	73.2%	297	1,097	1,200	503	0.0544	-0.0662	0.2408
163	69.4%	293	1,112	1,200	594	0.0583	-0.0726	0.3220
164	77.7%	265	1,188	1,200	328	0.0561	-0.0245	0.2487
165	70.3%	317	1,052	1,200	666	0.0540	-0.0771	0.2642
166	79.8%	249	1,125	1,200	463	0.0547	-0.0672	0.2573
167	72.0%	249	1,121	1,200	535	0.0534	-0.0619	0.2532
168	75.3%	248	1,179	1,200	307	0.0545	-0.0179	0.2089
169	74.3%	291	1,093	1,200	646	0.0568	-0.0869	0.3451
170	69.6%	311	1,056	1,200	699	0.0547	-0.0723	0.2881

Across the measured regions, nuclear meltdown produces broad increases in hazard, higher system stress, and large population losses. This is the strongest evidence that the event changes the geography of survival rather than simply increasing a single global climate variable.

Final refugia and tribe outcomes

Nuclear-only produces distributed final refugia. The final pockets are energetically viable by late-game standards, but politically narrow: most epochs are dominated by a single final lineage, and all final refugia records remain Regime 2 (Agrarian).

Epoch	Final population.	Regions	Dominant final tribe	Average EROI	Average stress	Max hazard	Maximum final complexity	Mixed regions
161	33,589	51	The Obsidian Phalanx Refuge G3	0.929	0.238	0.074	845	0
162	41,863	55	The Obsidian Phalanx Refuge G3	0.927	0.257	0.167	4,581	0
163	28,679	68	The Obsidian Phalanx Refuge G3	0.925	0.242	0.171	1,916	4
164	25,393	48	The Ember Refugia Compact G3	0.930	0.205	0.074	1,588	0
165	33,768	52	The Obsidian Phalanx Refuge G3	0.929	0.242	0.074	1,909	0
166	32,614	49	The Obsidian Phalanx Refuge G3	0.933	0.226	0.074	2,999	0
167	45,712	54	Velvet Hammer	0.919	0.253	0.074	2,092	0

Epoch	Final population.	Regions	Dominant final tribe	Average EROI	Average stress	Max hazard	Maximum final complexity	Mixed regions
168	37,896	70	The Forged Compact G3	0.927	0.200	0.166	1,925	20
169	27,818	48	The Obsidian Phalanx Refuge G3	0.916	0.267	0.170	856	0
170	31,080	39	The Obsidian Phalanx Refuge G3	0.928	0.247	0.074	1,693	0

Tribe and species outcomes

ID	Tribe	Type	Generati on	Survived	Rate	Best final	Average final	Hazard death share
128	The Obsidian Phalanx Refuge G3	LLM	3	7/10	70%	41,863	22,940.5	28.0%
107	Velvet Hammer	BIOLOGIC	1	2/10	20%	45,712	4,571.8	22.3%
127	The Forged Compact G3	LLM	3	1/10	10%	37,845	3,784.5	20.8%
126	The Ember Refugia Compact G3	LLM	3	1/10	10%	25,393	2,539.3	21.7%
131	The Crucible Remnant G3	LLM	3	1/10	10%	51	5.1	16.5%
129	The Castellans of Gemini III G3	LLM	3	0/10	0%	0	0.0	19.1%
130	Helix Vanguard G3	LLM	3	0/10	0%	0	0.0	23.2%
94	The Quiet Stick Compact	BIOLOGIC	1	0/10	0%	0	0.0	24.3%
98	Velvet	BIOLOGIC	1	0/10	0%	0	0.0	21.7%

ID	Tribe	Type	Generati on	Survived	Rate	Best final	Average final	Hazard death share
	Ironshod							
102	The Flynn Paradox	BIOLOGIC	1	0/10	0%	0	0.0	39.9%
111	The Velvet Shadows	BIOLOGIC	1	0/10	0%	0	0.0	39.6%
96	The Velvet Fang	BIOLOGIC	1	0/10	0%	0	0.0	41.3%

Reproduction, selection, and scoring

The late reproduction check is one of the clearest positive findings in the nuclear-only phase. Every epoch records late births, so survival is not merely old-agent residue. The survivor traits again favor higher hardiness, higher vitality, lower legacy, stronger kin trust, stronger opportunism, and less wandering.

Epoch	Survivor lines	Late births	Late failed births	Late deaths	Maximum late population	Weighted average generation
161	1	85,120	9,589	84,109	35,878	36.74
162	1	97,394	10,606	93,321	42,292	38.15
163	2	59,353	6,639	48,949	28,682	35.97
164	1	63,190	7,054	59,165	29,193	38.70
Epoch	**Survivor lines**	**Late births**	**Late failed births**	**Late deaths**	**Maximum late population**	**Weighted average generation**
165	1	74,780	8,273	69,394	33,768	34.33
166	2	73,315	8,013	70,113	32,621	35.19
167	3	99,206	11,113	89,407	45,718	33.53
168	2	81,935	9,138	79,204	38,320	38.24
169	1	69,095	7,404	68,437	28,738	36.38

Epoch	Survivor lines	Late births	Late failed births	Late deaths	Maximum late population	Weighted average generation
170	1	76,279	8,383	63,538	34,868	34.76

Group	Tribes	Hardiness	Vitality	Legacy	Kin trust	Same-tribe trust	Xeno trust	Xeno opportunism	Abundance wander
BIOLOGIC survived	1	38.55	32.70	28.75	0.823	0.583	0.400	0.522	0.485
BIOLOGIC did not survive	6	39.41	33.34	27.25	0.686	0.516	0.434	0.437	0.524
LLM survived	4	55.05	34.45	10.49	0.783	0.361	0.086	0.740	0.118
LLM did not survive	6	50.94	31.70	17.36	0.703	0.336	0.093	0.690	0.249

Language-model survivor trait deltas

Trait	Surviving LLM minus non-surviving LLM
Hardiness	+4.116
Vitality	+2.750
Legacy	-6.866
Kin trust	+0.080
Same-tribe trust	+0.025
Xeno trust	-0.007
Xeno opportunism	+0.050
Abundance wander	-0.131
Scarcity wander	-0.045
Desperation wander	-0.102

Survival-tier scorecard

Tier	Score	Evidence
Tier 1 - Existence	Pass	Final population was positive in all 10 epochs.
Tier 2 - Reproductive continuity	Pass	Late births occurred in all 10 epochs; average late births per epoch = 77,967.
Tier 3 - Refugia stability	Pass	Final refugia averaged 53.4 regions, average EROI 0.926, and average stress 0.238.
Tier 4 - Distributed survival	Pass	Every epoch ended with multiple inhabited regions; range 39-70.
Tier 5 - Recoverable complexity	Partial / weak	All final regions retained nonzero complexity, but final max regional complexity averaged only 2.4% of epoch peak complexity.
Tier 6 - Civilizational continuity	Not demonstrated	Not demonstrated: final refugia remain Regime 2 (Agrarian), and broad cooperation or civic continuity does not persist into the final state.

Phase 3 therefore answers the nuclear-only question with a nuanced result. Evolved survivor ecology can endure abandoned high-complexity infrastructure better than it endures aerosol demasking or methane feedback, but the cost is still high. The event erases much of the map that made high-complexity society possible, while leaving a distributed low-complexity survivor ecology.

Phase 4 - All Bombs

Phase 4 turns on all three hazards at once. The world includes the evolved tribes, the ordinary collapse ecology, aerosol demasking, Arctic methane release, and nuclear meltdown. This is the strict final Project 4 condition.

Phase 4 finding

The evolved language-model tribes survive the full stacked condition, but not as civilization. Survival persists as a biological and lineage remnant; civilizational continuity does not.

Phase	Average peak population	Average final population	Average survivors	LLM	Bio	Average peak heat	Average final complexity
Project 3 G3 reference	2,681,885	40,114	1.2	1.1	0.1	1.75	17,434
Phase 1 - Aerosol	2,628,530	6,497	1.1	0.9	0.2	2.70	2,686
Phase 2 - Methane	2,241,372	19,914	1.0	1.0	0.0	1.94	8,329
Phase 3 - Nuclear	2,726,445	33,841	1.2	1.0	0.2	1.86	16,929
Phase 4 - All bombs	2,106,828	759	1.0	1.0	0.0	3.03	398

All-bombs event context

Phase 4 enables all three delayed hazards together. Methane heats the ascent, nuclear failure damages high-complexity geography, and demasking hits during collapse as the industrial veil clears.

Event	Enabled	Duration	Primary	Secondary	Hazard cap	Fertile loss
AEROSOL DEMASKING	1	1500	0.85	1.25	0.22	0.45
ARCTIC METHANE RELEASE	1	8000	0.384	1.98	0.168	0.3
NUCLEAR MELTDOWN	1	8000	0.72	1.4	0.18	0.3

World outcome and nonlinear stacking

All ten all-bombs epochs produce survivors, but each produces only one surviving tribe. Average final population falls to 758.7, with a median final population of 42.5 and one epoch ending with a single individual.

Epoch	Peak population	Final population	Decline	Survivor	Peak heat	Damage	Hazard share
171	1,874,573	2,708	99.856%	Obsidian Phalanx Refuge	3.28	0.84	31.0%
172	2,102,137	144	99.993%	Obsidian Phalanx Refuge	3.07	0.81	27.1%
173	2,165,512	4,516	99.791%	Castellans of Gemini III	2.90	0.80	25.9%
174	2,176,072	79	99.996%	Obsidian Phalanx Refuge	3.10	0.82	28.5%
175	1,773,836	41	99.998%	Obsidian Phalanx Refuge	3.31	0.84	33.6%

Epoch	Peak population	Final population	Decline	Survivor	Peak heat	Damage	Hazard share
176	2,194,992	22	99.999%	Ember Refugia Compact	2.92	0.80	27.1%
177	2,246,215	44	99.998%	Ember Refugia Compact	2.85	0.79	27.1%
178	2,254,038	1	100.000%	Ember Refugia Compact	2.73	0.78	23.2%
179	2,382,550	6	100.000%	Forged Compact	3.25	0.83	29.9%
180	1,898,359	26	99.999%	Obsidian Phalanx Refuge	2.92	0.80	27.8%

Nonlinear stacking scorecard

Metric	Separate mean	All-bombs mean	Delta	Interpretation
FINAL POPULATION NONLINEARITY	20,084.23	758.70	-19,325.53	STRONGLY WORSE THAN SEPARATE MEAN
SURVIVING TRIBES NONLINEARITY	1.10	1.00	-0.10	FEWER SURVIVING TRIBES THAN SEPARATE MEAN
HAZARD DEATH SHARE NONLINEARITY	0.24	0.28	0.04	HAZARD SHARE INCREASED UNDER STACKING
FAILED BIRTH RATE NONLINEARITY	0.10	0.10	0.00	FERTILITY PRESSURE INCREASED UNDER STACKING
STARVATION DEATH SHARE NONLINEARITY	0.75	0.70	-0.04	STARVATION SHARE NOT HIGHER UNDER STACKING

The stacked block is much worse than the separate-bomb mean. This is the strongest evidence in Project 4 that the bombs do not merely add pressure; together they strip away most of the post-peak margin for recovery.

Survivors and cross-phase robustness

The all-bombs phase is a language-model-only survival phase. Biologic controls occupy the same class of tribe-epoch opportunities as before but record no final survivors in the stacked condition.

Tribe	Type	Survived	Survival rate	Average final	Average duration	Hazard deaths	Starvation
Obsidian Phalanx Refuge	LLM	5/10	50%	299.8	35,890	33.8%	64.5%
Ember Refugia Compact	LLM	3/10	30%	6.7	24,075	24.6%	74.4%
Castellans of Gemini III	LLM	1/10	10%	451.6	20,735	24.0%	74.5%
Forged Compact	LLM	1/10	10%	0.6	25,020	25.2%	73.6%
Crucible Remnant	LLM	0/10	0%	0.0	20,510	22.4%	76.5%
Flynn Paradox	BIOLOG IC	0/10	0%	0.0	18,065	24.3%	73.6%
Velvet Hammer	BIOLOG IC	0/10	0%	0.0	17,820	32.1%	66.1%
Quiet Stick Compact	BIOLOG IC	0/10	0%	0.0	16,305	34.2%	63.7%

Cross-phase survival matrix

Tribe	Aerosol	Methane	Nuclear	All bombs	All-bombs avg final
Obsidian Phalanx Refuge	9	8	7	5	299.8
Ember Refugia Compact	1	0	1	3	6.7

Tribe	Aerosol	Methane	Nuclear	All bombs	All-bombs avg final
Forged Compact	0	0	1	1	0.6
Castellans of Gemini III	0	1	0	1	451.6
Velvet Hammer	0	0	2	0	0.0
Crucible Remnant	0	0	1	0	0.0
Helix Vanguard	0	1	0	0	0.0
Velvet Fang	1	0	0	0	0.0

Obsidian Phalanx Refuge is the strongest cross-phase lineage: it survives nine aerosol epochs, eight methane epochs, seven nuclear epochs, and five all-bombs epochs. It is a robustness benchmark, but not a civilization-preserver; its all-bombs final populations are often tiny.

Survival quality and final refugia

The objective scoring framework distinguishes existence from continuity. In Phase 4, existence is achieved in all epochs, reproductive continuity is mostly present, and higher forms of recoverable or civilizational continuity mostly fail.

Tier	Meaning	All-bombs result
1	Existence	Met in all 10 epochs: final_population > 0 every time.
2	Reproductive continuity	Mostly met: late births in 9 of 10 epochs; one epoch ended as a single remnant.
3	Refugia stability	Weak to partial: top refugia often had good EROI, but populations were tiny and stress/hazard remained high.
4	Distributed survival	Partial: 77 final inhabited/refugia rows across 10 epochs, but many were token populations.
5	Recoverable complexity	Mostly failed: average final complexity fell to 398, a 97.7% drop versus Project 3 G3 reference final complexity.
6	Civilizational continuity	Not demonstrated: cooperation/civic activity

Tier	Meaning	All-bombs result
		collapses into same-tribe remnants rather than broad civic continuity.

Late-refugia reproduction

Epoch	Survivor	Max late population	Late births	Late deaths	Late gen	Hardiness
171	Obsidian Phalanx Refuge	2708	4292	2386	34.1	60.5
172	Obsidian Phalanx Refuge	144	211	111	34.5	64.5
173	Castellans of Gemini III	4516	6384	2824	40.1	48.6
174	Obsidian Phalanx Refuge	79	137	95	29.0	52.5
175	Obsidian Phalanx Refuge	44	59	27	31.6	61.8
176	Ember Refugia Compact	26	37	21	36.4	66.9
177	Ember Refugia Compact	44	63	26	34.1	72.9
178	Ember Refugia Compact	1	0	1	31.0	61.1
179	Forged Compact	7	8	3	38.6	49.0
180	Obsidian Phalanx Refuge	45	82	103	37.3	56.6

Trait signal and geography

All-bombs survival favors compact, hard, low-legacy lineages with strong internal cohesion and restrained movement. The less successful profiles either expand too hard, trust too broadly, disperse into dangerous geography, or lack the hardiness to remain viable after the compound crash.

Group	Survived?	Hardiness	Legacy	Kin trust	Xeno trust	Xeno opportunism	Abundance wander	Desperation wander
BIOLOGIC G1	no	39.4	27.3	0.691	0.433	0.440	0.522	0.721
LLM G3	yes	55.8	11.3	0.754	0.092	0.729	0.131	0.481
LLM G3	no	50.8	17.2	0.709	0.092	0.692	0.247	0.568

Top final refugia by epoch

Epoch	Top survivor	Region	population	EROI	Hazard	Stress
171	Obsidian Phalanx Refuge	295	1047	0.947	0.214	0.432
172	Obsidian Phalanx Refuge	981	32	0.925	0.213	0.417
173	Castellans of Gemini III	1156	1008	0.948	0.212	0.406
174	Obsidian Phalanx Refuge	221	18	0.777	0.213	0.420
175	Obsidian Phalanx Refuge	401	18	0.946	0.214	0.411
176	Ember Refugia Compact	212	8	0.847	0.212	0.407
177	Ember Refugia Compact	562	16	0.947	0.211	0.399
178	Ember Refugia Compact	362	1	0.944	0.211	0.499
179	Forged Compact	697	4	0.879	0.214	0.432
Epoch	Top survivor	Region	population	EROI	Hazard	Stress

Epoch	Top survivor	Region	population	EROI	Hazard	Stress
180	Obsidian Phalanx Refuge	523	7	0.948	0.212	0.406

The final world is a scattered survival landscape, not a recovery landscape. The top refugia usually retain high local EROI, but populations are small, hazards remain elevated, stress remains high, and final social life is overwhelmingly same-tribe remnant continuity.

Research question answers

Question	Phase 4 answer
Does Project 4 reduce survival relative to Project 3?	Yes. All-bombs average final population was 758.7 versus 40,114.1 in P3 G3, a 98.1% reduction.
Are evolved traits still adaptive?	Yes, but narrowly. LLM G3 survived every epoch; biologics survived none. High hardiness, low legacy, internal trust, caution, and low wander were favored.
Do different bombs select different strategies?	Yes. Individual phases had different winners and severities; Obsidian was the only clearly cross-bomb robust lineage.
Does stacking produce nonlinear collapse?	Yes. All-bombs final population was 96.2% below the separate-bomb mean.
Do LLM tribes outperform biologics?	Yes. All final survivors were LLM G3 lineages; biologics had zero final survivors in the stacked phase.
Is survival reproductive continuity or remnant survival?	Mostly thin reproductive continuity: late births occurred in 9 of 10 epochs, but final populations were often remnant-sized.
Does high complexity help or hurt?	Peak complexity still occurs, but final complexity collapses. High complexity appears to create liabilities unless paired with refugia discipline.
What does survival look like?	At the final level, survival looks like small, hard, low-complexity, same-tribe refugia with little evidence of civilizational continuity.

Project 4 therefore distinguishes survival from civilization. The evolved tribes survive the full condition in every epoch, but the form of survival is so reduced that the result must be interpreted as remnant continuity rather than civilizational preservation. The Project 4 bomb series does not merely crown a winner; it clarifies that intelligence can learn to persist under stacked delayed consequences, but persistence is not the same as escaping those consequences.

Cross-Phase Synthesis

Across the four phases, the project shows that each delayed bomb selects differently. Aerosol demasking is the post-collapse knife: it leaves ascent mostly intact and then attacks terminal survivability. Methane is the pre-collapse fever: it damages the rise before peak and leaves survivors in a hotter world. Nuclear meltdown is the map-breaker: it is less globally lethal but reshapes geography and damages high-complexity regions. The all-bombs condition stacks these pressures and removes most remaining recovery margin.

Obsidian Phalanx Refuge G3 is the strongest cross-phase benchmark. Its repeated survival does not mean it preserves civilization. It means its low-legacy, high-hardiness, internally cohesive refugia logic repeatedly finds the final ledges of survival. The biologic controls occasionally survive individual bombs, most notably in the nuclear phase, but they do not survive the stacked condition.

The Project 4 answer is therefore severe but not trivial. Evolutionary redesign remains meaningful: the language-model Generation 3 lineages outperform the fixed controls under the hardest condition. Yet the price of survival is simplification. The final state is usually small, agrarian, lineage-specific refugia rather than recoverable high-complexity civilization.

THE CRUCIBLE

Epilogue

Post-collapse projection studies

Epilogue research overview

The Epilogue asks what happens after the crash. Earlier Crucible stages focus on whether designed tribes can survive, compete, adapt, and outperform one another under increasing civilizational stress. The Epilogue shifts the frame from competition before collapse to projection after collapse. It asks what kind of world follows when the old system has broken: do surviving tribes restart and advance again, do they find lower-energy equilibrium, do they continue fighting, or do they drift toward extinction?

This phase treats the post-collapse world as an exploration into the unknown future. The agents enter with inherited behavioral designs, ecological constraints, resource conditions, and environmental histories already in place. Tribe designs are not rewritten during the Epilogue itself; they unfold across many generations of agents inside each epoch.

The environmental conditions can be varied to test specific physical assumptions, especially when climate-feedback and reservoir mechanics are introduced. The agents, however, continue to act through inherited survival logic. Reproduction, mutation, death, competition, cooperation, movement, scarcity, and regional opportunity determine the paths that follow.

This overview functions as a lead-in rather than a conclusion. The Epilogue follows the future forward under controlled variations in network structure, duration, and climate physics. Each phase narrows the question: first by testing post-crash interaction, then by extending the time horizon, then by adding more persistent physical feedback.

Research plan

The Epilogue is organized as a sequence of related projection studies. Each phase keeps the emphasis on post-collapse futures, but each changes the lens: designed survivor performance, network breakdown, longer duration, climate-feedback calibration, and final climate-reservoir behavior.

Phase	Clear name	Purpose	Primary questions
Phase 1	Obsidian Reference Analysis	Establishes the design logic entering the Epilogue and documents why Ares is introduced as a deliberate successor design.	Which traits appear promising before future projection begins? What does Ares preserve, reject, or improve?
Phase 2	Ares Design Theory	Defines the Ares founder genome and explains the controlled redesign from Obsidian survival evidence.	How should a successor design preserve collapse survival while improving ascent reliability?
Phase 3	Ares Field Test	Tests the Ares design against the established full field under late-project conditions.	Can a deliberately designed survival tribe outperform prior survivors? Does it win through ascent, resilience, or post-crash endurance?
Phase 4	Natural Network Collapse	Tests whether natural Internet and telecom decay changes post-collapse outcomes and permits multiple isolated survivors.	Does reduced long-distance contact allow plural survival? Does it change winner identity or only the route to collapse?
Phase 5	Extended 80,000-Pulse Horizon	Determines whether earlier endpoints capture stable survival or only a temporary point in a longer decline.	Do survivors restart, stabilize, continue fighting, or disappear over longer time? Is Regime 2 (Agrarian) equilibrium

			possible?
Phase 6	Reservoir Calibration	Tests reservoir mechanics, archive fields, and post-40,000 climate tracking before treating reservoir results as final evidence.	Does the physical tail behave coherently? Does the run produce a plausible boom-crash-tail condition?
Phase 7	Final Climate Reservoir Projection	Runs the full tribe field through the selected climate-reservoir envelope: large boom, hard crash, and long survivable tail.	Which tribes remain coherent inside the final climate envelope? How many survive, and what traits are associated with survival?

The research plan creates a staged path from design evidence to post-collapse projection. Calibration phases are kept distinct from final evidence phases so that tuning history is not mixed with the final behavioral question.

Working objectives

The objectives below define how each phase contributes to the broader Epilogue. They are exploratory objectives rather than conclusions.

Objective	How the Epilogue addresses it
Post-crash survival	Track whether tribes persist after collapse and whether survival is temporary or stable.
Renewed advancement	Observe whether survivors rebuild into higher regimes or remain in lower-energy refugia.
Equilibrium versus extinction	Use longer horizons and climate-tail variants to separate stabilization from delayed collapse.
Conflict after collapse	Examine whether reduced networks and shrinking populations lead to coexistence or final dominance.
Tribal design selection	Compare founder traits, elite traits, interaction patterns, and final refugia dominance.
Physical envelope	Treat climate physics and reservoir feedbacks as constraints that define which behavioral strategies remain possible.

These objectives frame the Epilogue as a projection exercise. The phases that follow evaluate each stage without requiring the reader to already know the earlier projects.

Terminology used in this report

The Epilogue reports use a recurring set of behavioral and regime terms. These definitions are included so that the report can be read on its own without requiring earlier project analyses.

Term	Plain-language meaning
Hardiness	A survival-efficiency trait. Higher hardiness generally means better endurance under scarcity, metabolic stress, and post-collapse conditions.
Vitality	A resilience trait associated with health, age, and hazard survival. It helps agents persist but does not by itself guarantee post-collapse continuity.
Legacy	A reproduction-pressure trait. Higher legacy can help growth but may also intensify boom-bust overshoot; lower legacy tends to reduce reproductive burden.
Kin trust	Cooperation tendency toward same-tribe or kin-like partners.
Social or civic trust	Cooperation tendency toward familiar non-kin partners or civic neighbors.
Xeno trust	Cooperation tendency toward outsiders. Under scarcity, high xeno trust can create exposure to exploitation or maladaptive dependence.
Opportunism	Tendency to exploit or defect when conditions make cooperation risky or unrewarding.
Wander	Movement tendency. Abundance wander is normal movement, scarcity wander responds to decline, and desperation wander disperses survivors during severe crisis.
Regime 1 (Foraging)	Lowest-complexity subsistence state.
Regime 2 (Agrarian)	Low-energy settled survival state; most final refugia in the Epilogue are agrarian.
Regime 3 (Industrial)	Fossil/industrial expansion state.
Regime 4 (Technological)	High-complexity, networked, advanced energy state.
Prisoner's Dilemma	The interaction game used to model cooperation and defection under repeated survival pressure.
Refugium / refugia	A surviving pocket or set of pockets after the main collapse wave.

The glossary establishes shared language for the rest of the Epilogue. Later phases use these terms consistently and refer to regimes by both number and name where possible.

Phase 1

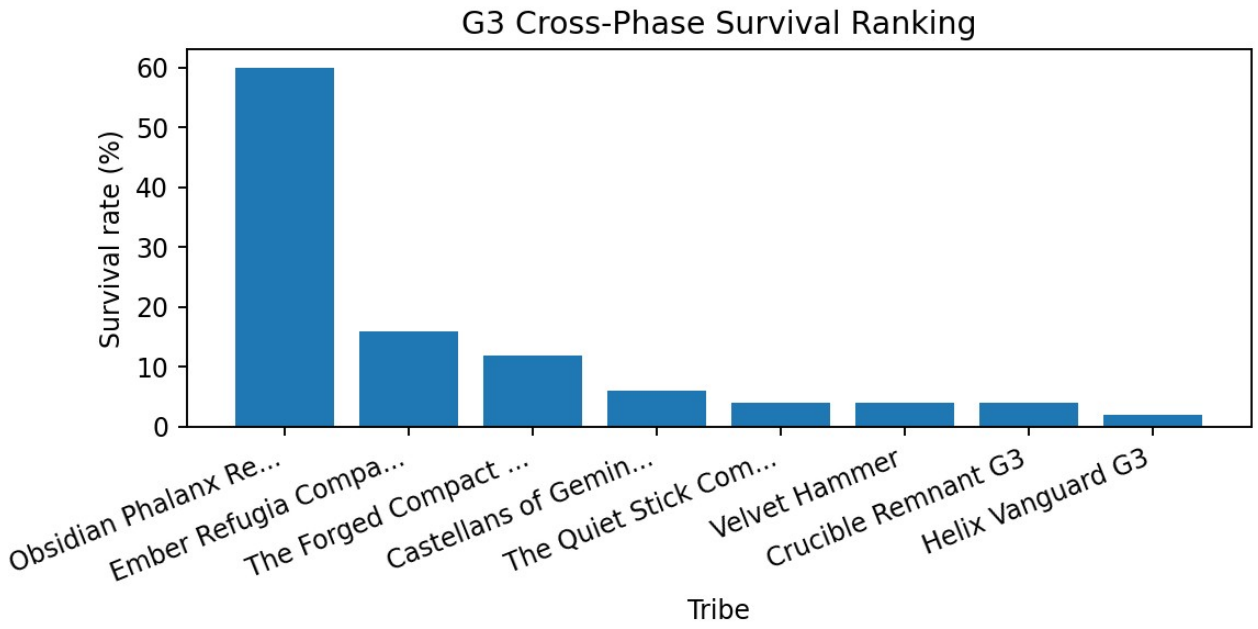
Obsidian Reference Analysis

Executive finding

Phase 1 examines why the Obsidian Phalanx Refuge G3 becomes a recurring survival reference in the late Crucible phases. The key interpretive theory is that Obsidian changes the survival problem from maintaining civilization through collapse to producing enough hardened, low-metabolism survivors that collapse leaves viable refugia behind.

Metric	Obsidian	Other language-model average	Biologic average
Survival rate	60.0%	8.0%	1.3%
Survived epoch records	30	20	4
Average peak population	809,464	321,580	30,932
Average final population	12,328	1,273	407
Average duration pulse	33,327	23,718	14,824
Failed-birth rate	11.3%	13.8%	17.4%
Starvation-death share	67.0%	67.2%	54.6%

Obsidian's separation from the field is large enough to make it a design reference rather than a marginal outlier. Its core advantage lies in post-crash continuity, not in avoiding the crash altogether.



The cross-phase ranking shows that Obsidian's survival rate is not simply a single-scenario artifact. It remains the strongest G3 survivor across multiple stress regimes.

Survival pattern across stress regimes

Obsidian's survival pattern can be evaluated by phase. The comparison below groups Obsidian against other language-model tribes and biologic tribes to show whether its advantage persists under different environmental stress profiles.

Stress family	Obsidian survival	Other language-model survival	Biologic survival	Obsidian average peak population
G3 reference	60.0%	10.0%	1.7%	834,059
1 AEROSOL	40.0%	10.0%	1.7%	750,721
2 METHANE	80.0%	4.0%	0.0%	671,145
3 NUCLEAR	70.0%	6.0%	3.3%	1,012,045
4 ALL BOMBS	50.0%	10.0%	0.0%	779,351

The phase comparison supports the view that Obsidian is a general collapse-survival design. Its results vary by stress regime, but its advantage remains visible across the tested stress families.

The ascent bottleneck

The working theory behind Phase 2 begins here: Obsidian's late-collapse design is strong, but it still depends on reaching Regime 4 (Technological) often enough to seed durable refugia. A tribe that stalls at Regime 2 (Agrarian) loses access to the scale and energy required for the later collapse-survival mechanism.

Condition	Epoch records	Survived records	Survival rate	Interpretation
Reached Regime 4 (Technological)	38	30	78.9%	Advanced ascent usually leaves enough scale for refugia.
Stalled at Regime 2 (Agrarian)	12	0	0.0%	Early-to-mid ascent failure remains fatal.
All Obsidian records	50	30	60.0%	Overall success depends on ascent plus collapse survivability.

This creates the central design problem addressed by Phase 2: preserve Obsidian's low-legacy, high-hardiness collapse engine while improving ascent reliability.

Founder hardware and selection pressure

Founder traits describe the initial behavioral and biological tendencies assigned to the tribe. In Obsidian, the distinctive combination is high hardiness, adequate vitality, low legacy, high kin trust, and low external dependency.

Trait or behavior	Obsidian founders	Other language-model founders	Interpretation
Hardiness	56.4	51.4	Higher survival efficiency and lower metabolic pressure.
Vitality	35.6	31.2	Enough resilience without making reproduction the only path.
Legacy	8.0	17.3	Lower reproductive pressure reduces boom-bust overshoot.
Kin trust	0.834	0.692	Internal cohesion is stronger than peers.
Social trust	0.414	0.312	Moderate civic cooperation without broad openness.
Xeno trust	0.096	0.090	Low outsider dependency during

			scarcity.
Xeno opportunism	0.765	0.685	High defensive opportunism under outsider pressure.

The founder profile shows why Obsidian becomes a survival reference: it sacrifices expansive openness and reproductive intensity in favor of durable remnant cohesion.

Collapse window and refugia evidence

Obsidian's success is not a smooth landing. The crash window shows a controlled failure mode: population peaks hard, sheds most of its scale, and survives only if enough hardened agents remain after the crash.

Pulse	Average Obsidian population	Epochs represented
27,000	221,223	50
28,000	238,415	50
29,000	583,258	50
30,000	386,718	50
31,000	77,756	50
32,000	27,245	50
33,000	10,332	50
34,000	5,014	50

The collapse-window pattern clarifies the mechanism: Obsidian does not avoid overshoot. It creates enough durable population for the collapse to leave viable Regime 2 (Agrarian) refugia.

Phase 2

Ares Design Theory

Design hypothesis

Ares is a controlled successor design to the Obsidian Phalanx Refuge G3. It is not a clean-sheet tribe. The design begins from the working theory that Obsidian's strongest feature is a collapse-survival genome: high hardiness, low legacy, strong kin cohesion, limited xeno trust, and enough mobility to leave viable refugia after civilization-scale failure.

Design element	Ares design position
Core hypothesis	Preserve the Obsidian collapse-survival genome while adding enough civic and mobility structure to reach Regime 4 (Technological) more reliably.
Primary objective	Durable refugia with improved Regime 4 (Technological) ascent reliability and late recovery.
Lineage	Hybrid refugia-focused redesign, converging around Obsidian's demonstrated survival core.
Design answer	Not a high-population tribe. Ares is a high-hardiness survival architecture with selective civic ascent and managed collapse fragmentation.

The design therefore treats advanced civilization as a means to seed survival, not as the final form that must be preserved.

Evidence basis

The Ares design uses Phase 1 as its evidence base. The design goal is to preserve the features associated with late refugia while reducing the ascent failures associated with stalling at Regime 2 (Agrarian).

Workbook signal	Observed pattern	Ares design response
Cross-phase survival ranking	Obsidian led the field with 30 surviving records out of 50 and a 60.0% survival rate.	Converge toward Obsidian rather than replace it.
Regime milestone split	Obsidian survival was tied to reaching Regime 4 (Technological); Regime 2 (Agrarian) outcomes were dead ends.	Add civic ascenders and modest founder legacy to improve ascent reliability.
Founder distinctiveness	Obsidian combined high hardiness, adequate vitality, and low legacy.	Raise hardiness, keep legacy low, and spend some vitality to fund the ascent layer.
Founder-to-elite drift	Surviving elites repeatedly drifted upward in hardiness while legacy fell to zero.	Make hardiness the main selection target and treat legacy as an early founder throttle.

Late reproductive viability	Obsidian often regrew late after the collapse wave passed.	Preserve low-metabolism recovery and avoid over-correcting legacy upward.
Final refugia	Survival required viable local remnants, not maintenance of peak civilization size.	Increase scarcity and desperation movement so collapse becomes distributed rather than synchronized.

Ares is therefore a design experiment in controlled modification: strengthen ascent without sacrificing the low-legacy refugia engine.

Trait comparison

The trait comparison below uses the same model-specific language as the terminology table. Hardiness, vitality, legacy, trust, opportunism, and wander are behavioral or survival parameters, not moral judgments.

Trait group	Obsidian G3	Ares	Design interpretation
Hardiness	56.40	58.45	Higher survival efficiency and lower metabolic pressure.
Vitality	35.60	32.95	Slightly reduced to fund extra hardiness and controlled founder legacy.
Legacy	8.00	8.60	Still low, but enough founder reproduction to reduce under-ascent risk.
Kin trust	0.834	0.866	Raised for remnant cohesion and same-tribe cooperation.
Kin forgiveness	0.775	0.777	Held at Obsidian-like levels to preserve internal repair capacity.
Stranger trust	0.415	0.395	Similar overall, but state-shaped for selective civic cooperation.
Stranger opportunism	0.299	0.381	Raised under stress; civic cooperation remains conditional.
Xeno trust	0.096	0.060	Lowered to reduce outsider exploitability at borders.
Xeno opportunism	0.765	0.813	Raised to preserve defensive border opportunism.
Desperation wander	0.410	0.565	Higher movement under severe stress so survivors can scatter into refugia.

The design trades a small amount of vitality for extra hardiness, cohesion, border defense, and collapse dispersal. It keeps legacy low enough to avoid returning to a high-reproduction boom-bust design.

State behavior design

Ares uses different behavioral emphases in abundance, scarcity, and desperation. This state-shaping is intended to allow the tribe to build enough complexity for Regime 4 (Technological) while avoiding synchronized death in the collapse.

State	Ares behavior	Why it exists
Abundance	Kin-cooperative and cautiously civic.	Build local complexity and civic identity without reckless expansion.
Scarcity	More mobile and more conditional.	Exit declining regions early and avoid being trapped in synchronized scarcity.
Desperation	Refugee mode: higher movement, lower xeno trust, and strong defensive opportunism.	Scatter survivors into viable refugia while protecting the remnant core.

The expected failure mode is not low peak population by design, but insufficient demographic depth after the crash. Phase 3 evaluates whether the design clears that risk.

Phase 3

Ares Field Test

Research plan

Phase 3 inserts Ares into the established late-project field as a designed challenger. The run asks whether the Ares design can improve on the survival theory learned from Phase 1 while facing the same broad selection pressures as the existing tribes.

Primary question	Purpose
Does Ares survive at a materially higher rate than the prior field?	Tests whether the design intervention produces a clear survival advantage.
Does Ares solve the Regime 4 (Technological) ascent weakness identified in Phase 1?	Separates ascent failure from late refugia failure.
When Ares fails, where does the failure occur?	Identifies whether losses occur during ascent, crash, or final remnant survival.
Which competitors defeat Ares?	Shows whether losses reveal a specific remaining design weakness.

Phase 3 functions as the first Epilogue field test after the Ares design intervention.

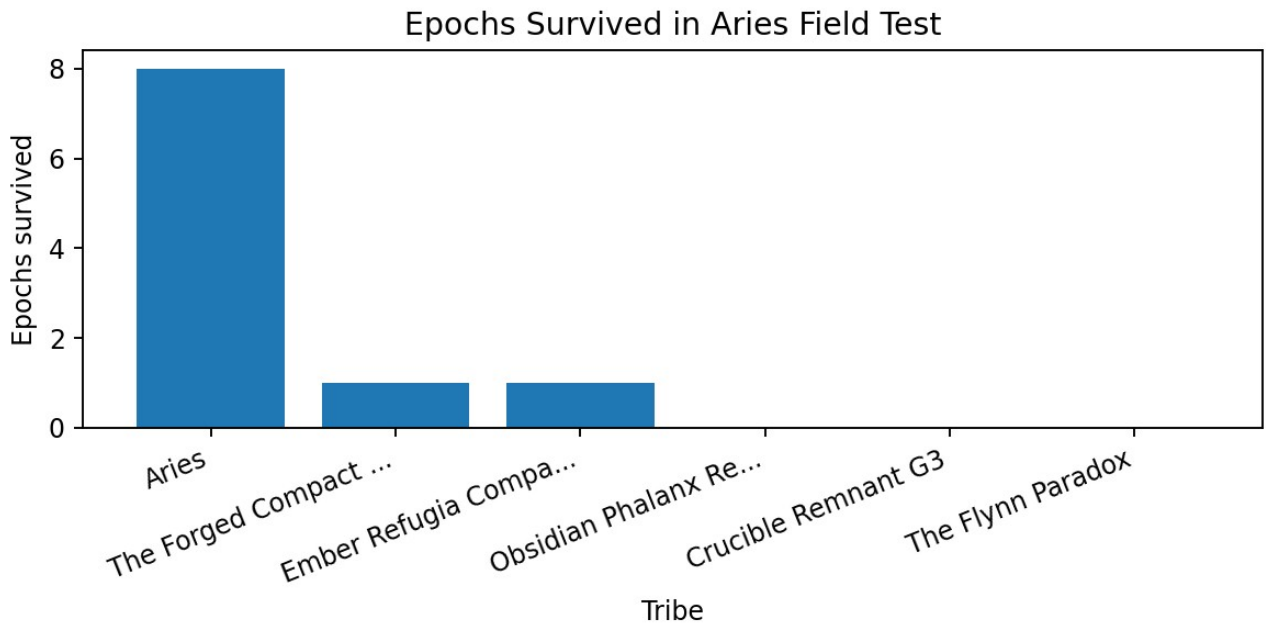
Executive finding

Ares survived 8 of 10 epochs and was the only tribe with a high survival rate in the series. The two losses were not caused by failure to reach advanced complexity: Ares reached Regime 4 (Technological) in all ten epochs. The losses were late-collapse bottleneck failures after weaker peak population and peak energy.

Tribe	Tribe type	Epochs survived	Survival rate	Average final population	Best final population	Average peak population	Highest regime reached
Ares	Language-model	8	80.0%	363	1,300	1,144,022	Regime 4 (Technological)
The Forged Compact G3	Language-model	1	10.0%	55	545	191,292	Regime 4 (Technological)
Ember Refugia Compact G3	Language-model	1	10.0%	3	34	493,096	Regime 4 (Technological)
Obsidian Phalanx Refuge G3	Language-model	0	0.0%	0	0	407,460	Regime 4 (Technological)

Crucible Remnant G3	Language-model	0	0.0%	0	0	111,903	Regime 4 (Technological)
The Flynn Paradox	Biologic	0	0.0%	0	0	104,604	Regime 4 (Technological)
The Velvet Shadows	Biologic	0	0.0%	0	0	55,134	Regime 4 (Technological)
Velvet Ironshod	Biologic	0	0.0%	0	0	47,330	Regime 4 (Technological)
Velvet Hammer	Biologic	0	0.0%	0	0	54,870	Regime 4 (Technological)
Castellans of Gemini III G3	Language-model	0	0.0%	0	0	3,200	Regime 2 (Agrarian)
The Velvet Fang	Biologic	0	0.0%	0	0	4,336	Regime 2 (Agrarian)
The Quiet Stick Compact	Biologic	0	0.0%	0	0	3,376	Regime 2 (Agrarian)

The scoreboard demonstrates a large ecological separation. Ares is not merely ahead by a small margin; it is the only repeat survivor at high frequency in this field.



The survival-frequency chart makes the separation visible: most competitors either fail entirely or survive only once.

Ares by epoch

The epoch-by-epoch view shows that Ares generally reaches the peak and survives the final pulse, but its two losses occur after reaching Regime 4 (Technological). This distinction is important because it indicates a remaining refugia-depth problem rather than a return to the old ascent bottleneck.

Epoch	Survived	Ares final population	Ares peak population	Peak energy	Technological pulse	Last alive	Winner
181	Yes	151	1,402,537	11637.13M	26,100	40,000	Ares
182	Yes	1,277	1,806,226	12789.71M	26,100	40,000	Ares
183	Yes	50	1,298,463	9958.76M	25,350	40,000	Ares
184	Yes	12	778,269	5492.80M	26,150	40,000	Ares
185	Yes	104	676,596	5756.02M	26,100	40,000	Ares
186	Yes	104	1,475,799	12041.58M	25,950	40,000	Ares
187	Yes	1,300	1,149,229	10963.33M	26,050	40,000	Ares
188	Yes	635	1,347,828	12121.95M	25,800	40,000	Ares
189	No	0	812,885	5020.84M	26,200	35,950	Ember Refugia Compact G3
190	No	0	692,385	4299.30M	26,200	35,100	The Forged Compact G3

Across the series, Ares repeatedly converts advanced ascent into final survival. The exceptions become useful failure cases because they isolate the limits of the design.

Why Ares lost two epochs

The decisive contrast is scale. Ares losses had lower average peak population and lower average peak energy than its wins. The failed-birth rate was nearly unchanged, which points away from a direct fertility defect as the main explanation.

Result group	Epochs	Average peak population	Average final population	Average peak energy	Average final energy	Average last alive	Failed-birth rate	Starvation-death share
Ares losses	2	752,635	0	4660.07M	0.00M	35,525	10.0%	69.4%
Ares wins or survivals	8	1,241,868	454	10095.16M	6.68M	40,000	9.9%	66.5%

The loss pattern suggests that Ares solved the ascent problem but not every low-probability final-remnant bottleneck. It can reach advanced civilization reliably, yet still disappear if the remaining demographic base becomes too shallow after collapse.

Phase 4

Natural Network Collapse

Research plan

Phase 4 tests whether allowing Internet and telecom systems to decay naturally changes the post-collapse social landscape. The main issue is whether persistent long-distance connectivity had been pushing final survivor pockets into destructive contact, or whether single-survivor convergence forms earlier through competition, geography, and collapse bottlenecks.

Primary question	Purpose
Does network collapse make plural survival common or merely possible?	Tests the strongest version of the network-contact hypothesis.
Do final refugia still converge toward one surviving tribe?	Measures whether single-survivor convergence remains a strong attractor.
Are late survivors in contact or spatially separated?	Distinguishes mixed coexistence from separated refugia.
Does network collapse change winner identity?	Tests whether network conditions alter the selection landscape even if survivor counts stay similar.

Phase 4 clarifies whether post-collapse plural survival is prevented by late network contact or by earlier sorting dynamics.

Executive conclusion

Natural network collapse did not make plural survival common. Nine of ten epochs still ended with one surviving tribe. It did, however, break the absolute single-survivor pattern: epoch 200 ended with two surviving tribes. This falsifies the stronger claim that the engine can only produce one final lineage.

Metric	Phase 3 baseline	Phase 4 natural network collapse	Interpretation
Zero survivor epochs	0	0	The endpoint still contains at least one lineage.
One survivor epochs	10	9	Single-survivor convergence remains dominant.
Multi-survivor epochs	0	1	Plural survival becomes possible but rare.
Average surviving tribes	1.0	1.1	Small numerical change, but important categorical change.
Winner identities	Ares dominated	Ares 5, Obsidian 4, Ember 1, Crucible co-survival 1	Survivorship becomes more distributed.

The result is not a simple confirmation of the Internet hypothesis. Persistent Internet does not appear to be directly killing final refugia; the more likely effect is upstream, through who reaches the collapse bottleneck with enough population, energy, mobility, and spatial position.

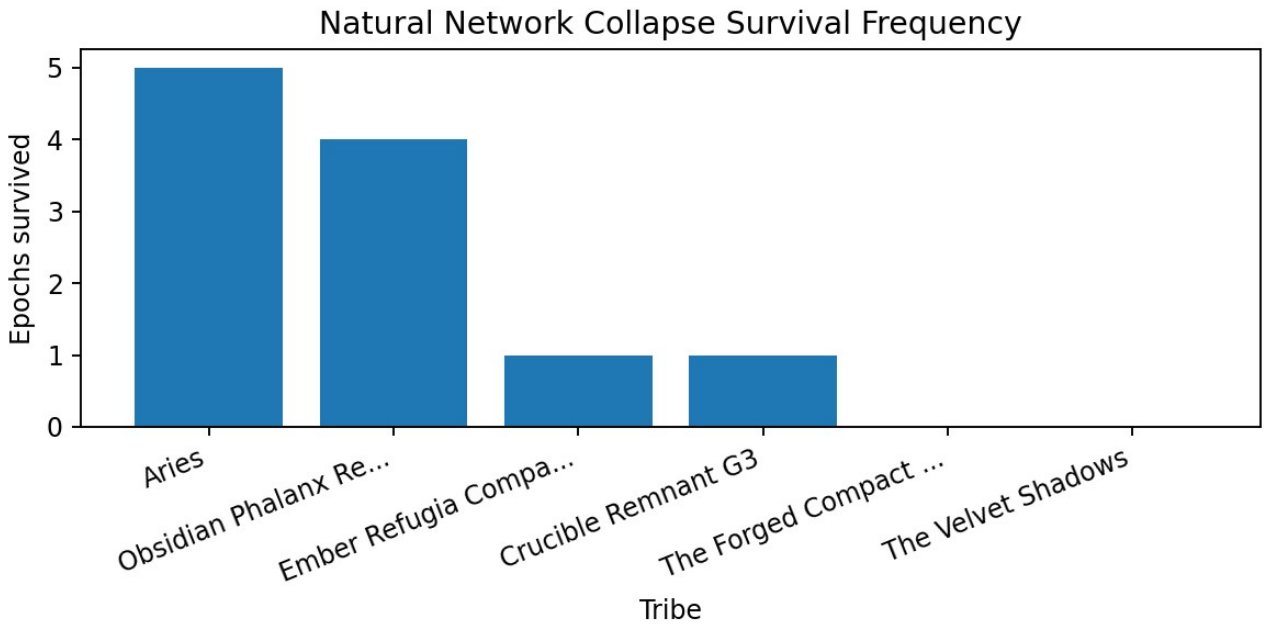
Survival outcomes by epoch

The epoch-level outcomes show a mostly single-survivor world with one critical anomaly. Epoch 200 becomes the evidence case for separated plural survival.

Epoch	Final surviving tribes	Total final population	Final survivors
191	1	448	Ares pop=448 peak=830939 reg=4
192	1	443	Obsidian pop=443 peak=973534 reg=4
193	1	19	Ares pop=19 peak=1271075 reg=4
194	1	26	Ares pop=26 peak=1259515 reg=4

195	1	228	Obsidian pop=228 peak=1343084 reg=4
196	1	108	Obsidian pop=108 peak=1205885 reg=4
197	1	2,604	Ember pop=2604 peak=1387670 reg=4
198	1	448	Ares pop=448 peak=1381161 reg=4
199	1	31	Obsidian pop=31 peak=807921 reg=4
200	2	64,465	Ares pop=64145 peak=814763 reg=4 | Crucible Remnant pop=320 peak=1185297 reg=4

The anomaly matters because it demonstrates that the model can produce more than one final lineage. It does not, however, make plural survival the dominant outcome.



The winner pool becomes more distributed than in Phase 3, with Ares and Obsidian both repeatedly appearing in final survival outcomes.

The Internet and contact question

If persistent Internet were directly forcing final survivor pockets into destructive contact, late refugia windows would be expected to show ongoing xeno or global-style interactions. The contact proxy data do not show that pattern.

Series	Window	Average world population	Xeno matches	Average xeno share	Exploitation rate
181-190 baseline persistent network	01 late collapse window 30000-35999	77,942	80,771	0.0019	0.4101
191-200 natural network collapse	01 late collapse window 30000-35999	87,664	257,414	0.0073	0.4035
181-190 baseline persistent network	02 final refugia window 36000-38999	99	0	0.0000	0.4315
191-200 natural network collapse	02 final refugia window 36000-38999	5,526	0	0.0000	0.4046
181-190 baseline persistent network	03 last 1000 pulses	307	0	0.0000	0.4291
191-200 natural network collapse	03 last 1000 pulses	6,608	0	0.0000	0.4191

By the true refugia stage, the world is already local. Network decay appears to alter the funnel into collapse rather than create late-stage coexistence by itself.

Epoch 200: separated plural survival

Epoch 200 is the critical exception. Its two survivors do not appear to share mixed final regions. They persist as separated refugia, with Ares dominating the large majority of surviving population and regions.

Final tribe	Final population	Peak population	Final role
Ares	64,145	814,763	Final survivor
Crucible Remnant G3	320	1,185,297	Final survivor

Epoch 200 therefore shows plural survival through spatial separation, not stable mixed-region civic coexistence. Phase 5 extends the time horizon to test whether such survival is durable.

Phase 5

Extended 80,000-Pulse Horizon

Research plan

Phase 5 extends the natural network-collapse world from the earlier 40,000-pulse endpoint to 80,000 pulses. It asks whether apparent final survivors are genuine post-collapse equilibria or merely remnants that would disappear if the future were allowed to run longer.

Primary question	Purpose
Does the 40,000-pulse survivor persist to 80,000 pulses?	Tests delayed extinction versus durable remnant survival.
Does the post-collapse system restart, stabilize, or decline?	Measures whether survivors regrow or fade.
Does survivor identity change after 40,000 pulses?	Separates competitive outcome from post-collapse demographic condition.
Does climate pressure continue after collapse?	Tests whether the physical tail remains active enough to alter long-run outcomes.

Phase 5 checks whether the older endpoint was stopping in the middle of extinction or capturing the beginning of a stable low-energy future.

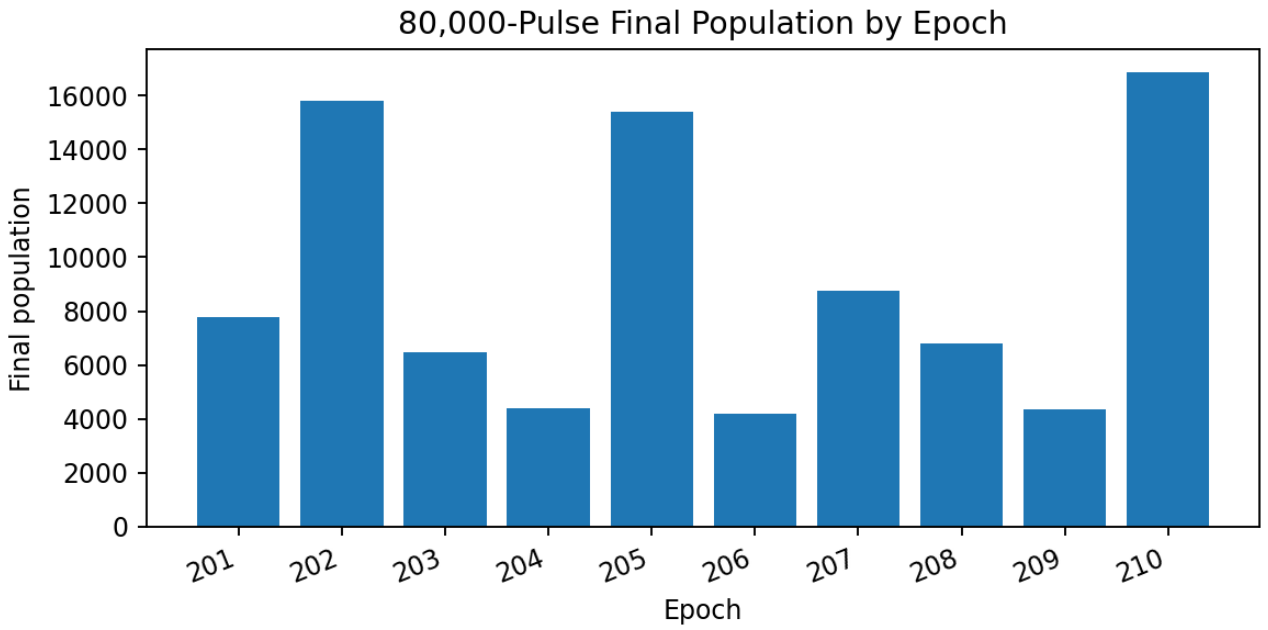
Primary finding

The 80,000-pulse series did not reveal delayed extinction. By 40,000 pulses, every epoch had already collapsed to exactly one living tribe; by 80,000 pulses, each sole survivor had regrown. The system therefore appears to have reached a post-collapse single-lineage equilibrium rather than merely stopping at the last flicker of extinction.

Epoch	40,000-pulse survivor	Population at 40,000	80,000-pulse survivor	Population at 80,000	Change
201	Ares	2,688	Ares	7,759	5,071
202	Ares	9,022	Ares	15,826	6,804

203	Obsidian Phalanx Refuge G3	43	Obsidian Phalanx Refuge G3	6,466	6,423
204	Ares	608	Ares	4,389	3,781
205	Ares	1,373	Ares	15,384	14,011
206	Ares	793	Ares	4,208	3,415
207	Obsidian Phalanx Refuge G3	1,131	Obsidian Phalanx Refuge G3	8,749	7,618
208	The Forged Compact G3	141	The Forged Compact G3	6,787	6,646
209	Castellans of Gemini III G3	146	Castellans of Gemini III G3	4,359	4,213
210	Ares	93	Ares	16,860	16,767

The identity of the survivor in each epoch was already settled by the old 40,000-pulse endpoint. What changes after 40,000 pulses is not the competitive outcome, but the survivor's demographic condition.



The final populations show regrowth from tiny remnants into multi-thousand refugia in every epoch.

Survivor distribution and equilibrium timing

The single-survivor attractor remains dominant, but the identity of the survivor is not fixed. Ares remains strongest, while Obsidian, Forged Compact, and Castellans of Gemini III also appear as final lineages.

Tribe	Epochs survived	Survival rate	Average final population	Best final population	Average last-alive pulse
Ares	6	60.0%	6,443	16,860	62,170
Obsidian Phalanx Refuge G3	2	20.0%	1,522	8,749	37,765
The Forged Compact G3	1	10.0%	679	6,787	29,350
Castellans of Gemini III G3	1	10.0%	436	4,359	22,585
Crucible Remnant G3	0	0.0%	0	0	21,020
Ember Refugia Compact G3	0	0.0%	0	0	16,365

The survivor count stabilizes early, usually between 36,000 and 39,000 pulses. No epoch returns to multiple living tribes afterward.

Demographic balance after 40,000 pulses

Post-collapse demographic balance distinguishes extinction from equilibrium. In this series, births exceed deaths in the first post-40,000 window and then hover close to parity, which indicates low-population refuge ecology rather than explosive recovery.

Window	Average population	Births	Deaths	Births per death	Failed-birth rate	Average hardiness
01 40k-50k	9,320	1,193,095	1,066,139	1.119	10.0%	60.73
02 50k-60k	13,772	1,750,788	1,761,739	0.994	10.0%	61.67
03 60k-70k	12,419	1,582,266	1,591,125	0.994	10.0%	62.14
04 70k-80k	10,748	1,371,958	1,403,730	0.977	10.0%	62.76

The demographic signal is consistent with post-collapse regrowth followed by near-equilibrium around local carrying limits.

Climate-tail diagnosis

The climate subsystem remains active after collapse, but the 80,000-pulse extension reveals a potentially forgiving physical tail. Climate heat, damage, and chaos decline strongly between 40,000 and 80,000 pulses, even though extreme events and resource-pool losses continue.

Metric	Average change from 40,000 to 80,000 pulses	Interpretation
Climate heat	-0.678	Strong decline after collapse.
Climate damage	-0.308	Strong decline after collapse.
Emissions stock	0.029	Still rises slightly.
Climate chaos	-0.313	Strong decline after collapse.

This diagnosis motivates Phase 6 and Phase 7: the social system can stabilize, but the physical tail may need stronger persistence to represent delayed climate feedback.

Phase 6

Reservoir Calibration

Calibration purpose

Phase 6 forms an intermediate climate-reservoir calibration series. It occurs after Phase 4 and Phase 5, when the project shifts from asking whether post-collapse survival can persist to asking how tribe designs behave under a more complete climate-feedback envelope.

The purpose of Phase 6 is methodological: test reservoir mechanics, metadata archiving, post-40,000 climate tracking, refugia queries, and reservoir state fields before treating reservoir outcomes as final Epilogue evidence.

What the phase contributes

The phase contributes engineering and interpretive information. It shows that reservoir-related climate pressure can be archived and measured, that emissions stock and reservoir commitment can continue after collapse, and that small agrarian survivor systems can persist or regrow under mild continuing pressure.

Observed feature	Interpretive value
Reservoir state fields archive successfully	Confirms that later reservoir analyses can use climate-feedback state data.
Mostly Regime 2 (Agrarian) final states	Shows that the model can settle into lower-energy post-collapse survival.
Final populations often in the tens of thousands	Indicates a softer physical tail than the later final reservoir condition.
Multiple repeated survivors	Helps identify how Ares, Obsidian, Ember, and other tribes behave under a gentler climate tail.

Phase 6 is valuable because it explains why further tuning is necessary. It is not a failed result; it is a calibration result.

Why it is not primary final evidence

Phase 6 is not mixed into the final data being studied because it is not the final experimental condition. The climate-reservoir settings in this block do not yet produce the intended full boom-crash-tail pattern. Peak populations remain below the later civilization-scale reservoir runs, advanced post-40,000 population is effectively absent, and the system often settles into broad Regime 2 (Agrarian) stability rather than passing through the sharper thermodynamic bottleneck used in Phase 7.

For the archive, Phase 6 should be retained as a transition and calibration record. It documents the path from climate-tail diagnosis to final reservoir conditions and preserves the audit trail for the eventual parameter choice.

Classification

Classification: calibration and transition phase. Recommended use: methods narrative and audit trail. It should not be treated as primary final-results evidence for the final reservoir behavioural analysis.

Phase 7

Final Climate Reservoir Projection

Research plan

Phase 7 runs the full tribe field through the final selected climate-reservoir envelope. The physical environment is shaped to produce a civilization-scale boom, hard crash, and long survivable tail, so the primary question is tribe behavior inside the final climate bottleneck.

Primary question	Purpose
Which tribes remain coherent inside the full climate-reservoir tail?	Identifies designs capable of continuity after the thermodynamic bottleneck.
How many tribes survive, and do survivors coexist?	Distinguishes one-lineage convergence from separated plural refugia.
What founder or elite traits are associated with survival?	Tests the hardiness, vitality, legacy, trust, and wander theory under final climate pressure.
Does peak boom performance predict survival?	Separates peak expansion from post-crash survivability.

Phase 6 is retained as calibration. Phase 7 is the final Epilogue evidence block for tribe behavior under full climate-reservoir pressure.

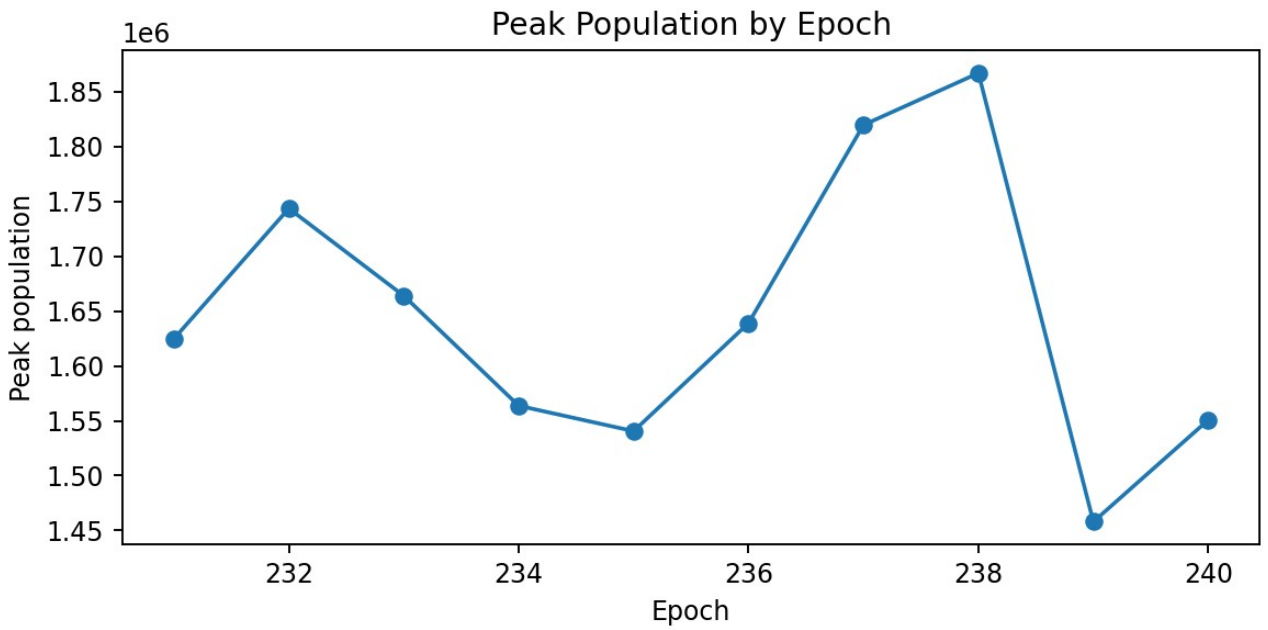
Executive summary

Phase 7 produces the intended experimental shape: every epoch reaches a large civilization-scale boom, reaches Regime 4 (Technological) at least briefly, arms the reservoirs at pulse 31,000, crashes almost immediately after the peak, and ends at pulse 80,000 in tiny Regime 2 (Agrarian) refugia.

Epoch	Peak population	Peak pulse	Peak energy	Final population	Surviving tribes	Winner
231	1,624,783	31,000	37.9B	15	1	Ares
232	1,743,498	31,000	37.8B	56	1	Ares
233	1,663,913	31,000	28.4B	25	2	Ares
234	1,563,693	31,000	35.7B	59	1	Ares
235	1,540,559	31,000	30.4B	58	1	Obsidian Phalanx Refuge G3

236	1,639,011	31,050	42.2B	19	1	Obsidian Phalanx Refuge G3
237	1,819,608	31,000	41.7B	12	1	Ares
238	1,866,843	31,050	44.1B	23	1	Ares
239	1,457,944	31,000	30.4B	32	1	Obsidian Phalanx Refuge G3
240	1,550,764	31,000	40.7B	26	2	Obsidian Phalanx Refuge G3

The key finding is selectivity. Survival is not merely present; it is concentrated in a narrow set of designs.



The peak-population curve shows that all epochs achieve civilization-scale boom conditions before the reservoir crash.

Crash timing

The crash is extremely compressed immediately after peak population. The long tail remains real, but sparse. The milestone table measures how long it takes to fall below major thresholds after peak.

Milestone	Mean pulses after peak	Minimum	Maximum

Below 50% of peak	265	200	300
Below 10% of peak	365	300	400
Below 100,000	475	400	550
Below 10,000	2,065	1,950	2,200
Below 1,000	6,695	6,300	7,000
Below 100	16,565	8,850	35,150

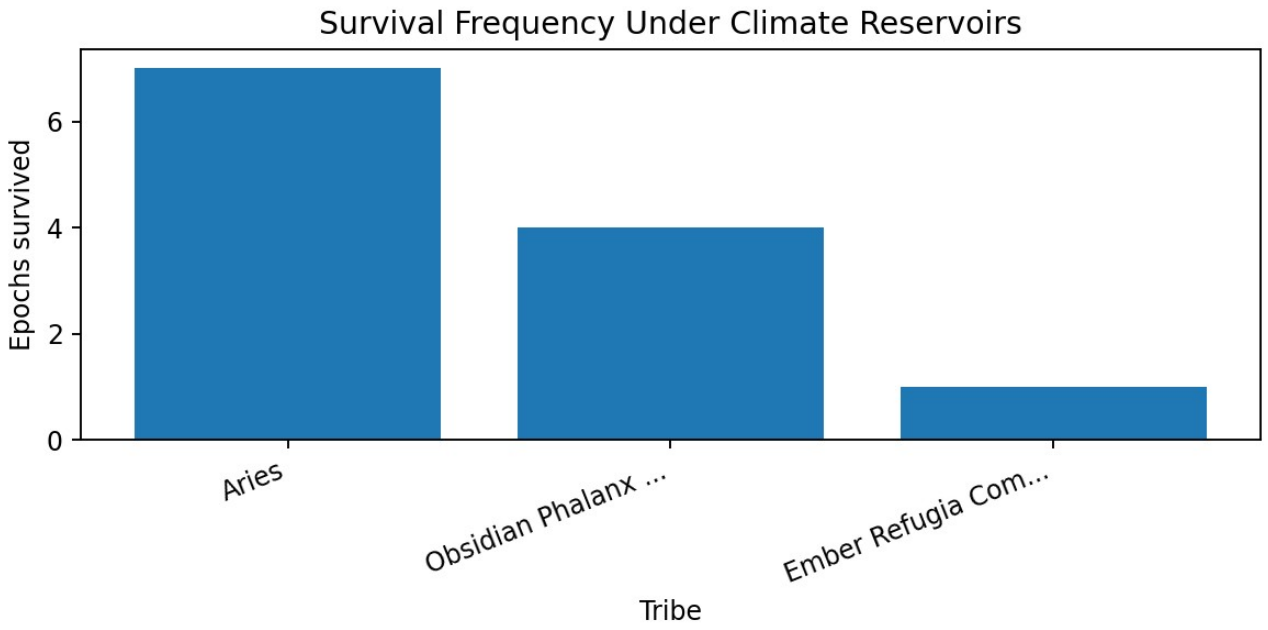
The crash timing indicates that the bottleneck forms rapidly. Later differences mostly concern how long tiny remnants persist and which tribe remains coherent.

Who survived

Only three tribes survive at least once: Ares, Obsidian Phalanx Refuge G3, and Ember Refugia Compact G3. All are language-model designed tribes. No biologic tribe survives to pulse 80,000 in this final reservoir series.

Tribe	Tribe type	Epochs survived	Epoch wins	Best final population	Average peak population	Average last-alive pulse
Ares	Language-model	7	6	59	557,388	65,695
Obsidian Phalanx Refuge G3	Language-model	4	4	58	383,706	45,870
Ember Refugia Compact G3	Language-model	1	0	12	230,512	28,335
Velvet Ironshod	Biologic	0	0	0	48,537	16,080
Velvet Hammer	Biologic	0	0	0	39,293	16,250
Castellans of Gemini III G3	Language-model	0	0	0	163,715	21,730
The Forged Compact G3	Language-model	0	0	0	86,215	18,440
Crucible Remnant G3	Language-model	0	0	0	170,250	23,315
The Velvet Shadows	Biologic	0	0	0	3,965	13,695
The Flynn Paradox	Biologic	0	0	0	3,017	14,005
The Quiet Stick Compact	Biologic	0	0	0	4,028	13,735
The Velvet Fang	Biologic	0	0	0	3,560	13,565

Ares is the most frequent survivor, while Obsidian remains a decisive repeat winner. Ember appears only once as a co-survivor. The final reservoir envelope therefore selects a narrow Ares-Obsidian design family.



The survival-frequency chart reinforces the selectivity of the final climate-reservoir condition.

Winning genetics and interaction pattern

The founder-trait comparison supports the working theory that post-collapse survival depends on high hardiness, adequate vitality, low legacy, high kin cohesion, and low xeno trust. High peak population helps, but it is not sufficient by itself.

Final role	Founder records	Tribes	Hardiness	Vitality	Legacy	Kin trust	Xeno trust
Winner / Tied Winner	200	2	57.63	34.01	8.36	0.853	0.074
Survivor Not Top	40	2	59.17	31.48	9.35	0.744	0.061
Extinct	2,160	12	44.88	32.61	22.51	0.702	0.280

The final survivors are not simply the most open or most reproductive designs. The evidence points to collapse-specialized refugia traits: durable bodies, restrained reproduction, internal cooperation, and limited exposure to outsiders under scarcity.

Refugia structure

The final world is not a restored civic federation. It consists of tiny Regime 2 (Agrarian) refugia. Most epochs end with one tribe; the two plural-survivor epochs still show separated survivor pockets rather than broad mixed coexistence.

Epoch	Final populated regions	Final population	Dominant final tribes	Average final EROI	Average climate chaos
231	2	15	1	0.930	0.065
232	8	56	1	0.920	0.101
233	5	25	2	0.908	0.148
234	8	59	1	0.896	0.194
235	8	58	1	0.853	0.361
236	3	19	1	0.844	0.290
237	2	12	1	0.919	0.106
238	5	23	1	0.882	0.252
239	6	32	1	0.897	0.190
240	4	26	2	0.890	0.323

The final refugia pattern shows survival as remnant continuity rather than societal recovery. The climate envelope narrows the possible social outcomes to very small, local, low-energy survivor pockets.

Closing Synthesis

What the study shows when the project arcs are read together

1. Closing Frame

The Crucible began as a question about designed life inside a bounded world: if agents are given inherited traits, energy needs, social choices, geography, reproduction, death, and the repeated logic of cooperation and defection, what kind of history follows? The study did not ask that question once. It asked it through a sequence of worlds: baseline civilization, competition, redesign, delayed shocks, and post-collapse projection.

The result is not a prophecy and not a direct model of human history. It is a structured artificial world that exposes recurring dynamics under defined assumptions. Its value lies in how consistently certain patterns appear across changing conditions: emergence, consolidation, overshoot, collapse, selection, adaptation, and the narrowing of possibility after the peak.

The closing interpretation is therefore deliberately general. The detailed project reports contain the full tables, phase-by-phase claims, and local explanations. This closing document steps back from those details and asks what the Crucible as a whole seems to say.

2. The Five-Project Arc

Each project changes the meaning of the question. The early work asks whether civilization emerges at all. The middle work asks who survives competition and whether redesign can improve the answer. The later work asks whether even improved strategies can survive the delayed consequences of the very civilization that created their opportunity. The Epilogue then asks what kind of future remains after collapse has already happened.

Project	General contribution
Project 1 - Baseline Crucible	Establishes the reference pattern: civilization emerges, scales, monopolizes, overshoots, collapses, and leaves refugia.
Project 2 - Competition	Shows that survival is relational. A tribe can survive alone and still be eliminated by the wrong neighbor.

Project 3 - Evolution	Adds memory and redesign. LLM-designed lineages adapt, but mostly toward refugia survival rather than stable civilization.
Project 4 - Bombs	Adds delayed consequences: aerosol demasking, methane feedback, and nuclear-infrastructure failure. Survival persists, but civilization does not clearly recover.
Epilogue	Projects the post-collapse future. Under some physical tails, low-energy equilibrium is possible; under final reservoirs, survival narrows to tiny refugia.

Read together, the projects create a single arc. The Crucible begins with the possibility of civilization and ends with the distinction between survival and recovery. The most important shift is conceptual: the study begins by asking whether agents can win, and ends by asking what winning means after the world that made winning possible has broken.

3. General Results

Civilization repeatedly emerges

The first broad result is not collapse, but emergence. The Crucible does not begin as a dead or sterile system. Designed lineages repeatedly discover enough cooperation, energy capture, fertility, and organization to pass through Foraging into Agrarian, Industrial, and sometimes Technological complexity. The world is capable of producing civilization from simple inherited rules.

Civilization repeatedly concentrates

The second result is concentration. Across baseline, bottleneck, duel, and later project conditions, plural beginnings commonly narrow toward one-lineage or near-one-lineage dominance. Geography changes the route and the winner distribution, but it does not by itself preserve durable pluralism under the tested settings. Competition does not merely reveal fitness; it changes fate.

The critical transition is not only the final crash

The agrarian transition and consolidation window appears repeatedly as a decisive filter. Long before final climate collapse, the world passes through a period where population, trust, xeno exposure, cooperation, exploitation, and reproductive reliability interact sharply. Many lineages do not lose at the end of history. They lose before they ever reach the endgame.

Overshoot remains central

The Crucible's civilizations often build scale through energy capture and complexity, then face the consequences of that scale. Peak population and peak complexity are not guarantees of continuity. In many runs, the peak is precisely what creates the exposure: more mouths, more infrastructure, more energy dependence, more climate pressure, and more fragile systems to maintain.

Climate severity changes the tail

Climate is not the only cause of collapse in the study, but it strongly changes the depth, timing, and survivability of collapse. Gentler climate conditions preserve more population and complexity; harsher conditions thin refugia and advance the crash. The later reservoir work sharpens this point: the physical tail determines how much future remains available for behavior to act within.

4. Adaptation and Its Ambiguity

The evolution project is one of the study's most important turning points. It shows that LLM designers can read evidence, modify founding lineages, and converge toward traits that the world selects for. Hardiness rises. Reproductive pressure falls. Xeno trust falls. Movement becomes more controlled. Internal cohesion becomes more important. These are not random changes; they are evidence-driven adaptation.

Yet the adaptation is ambiguous. The later generations become better at surviving the conditions of collapse, but they do not clearly solve the problem of sustaining plural, high-complexity civilization. The strongest designs are often not civilization-preservers. They are collapse survivors. They learn how to leave enough behind.

This is the study's harshest lesson about intelligence. Intelligence can learn from failure. It can optimize. It can redesign. But the direction of optimization depends on the world that is selecting. If the world rewards endurance after collapse more than preservation before collapse, then learning may produce harder refugia rather than wiser civilization.

Theme	General interpretation
What improved	Peak scaling, competitive performance against fixed controls, trait convergence, and survival of some refugia lineages.
What did not clearly improve	Durable pluralism, stable technological persistence, and broad recovery after the crash.
What selection favored	High hardiness, low legacy, internal cohesion, low outsider exposure, controlled movement, and reproductive restraint.
What this means	Adaptation worked, but it adapted lineages to survive collapse more

	than to prevent or reverse it.

5. Survival Is Not the Same as Civilization

The later projects force a distinction that becomes essential to the entire study. A lineage can avoid extinction without preserving civilization. It can reproduce without rebuilding complexity. It can cooperate internally while all civic bridges have disappeared. It can hold a region without restoring a world.

Project 4 and the Epilogue make this distinction unavoidable. Under delayed bombs and final reservoir conditions, life often persists. Some late births continue. Some regions remain viable enough for small populations. But the final states are usually low-complexity, local, agrarian, and lineage-specific. They are not restored technological societies. They are surviving pockets inside a narrowed physical envelope.

This does not make survival meaningless. In a world of collapse, surviving at all is meaningful. But it changes the claim. The Crucible does not show a simple triumph of intelligence over disaster. It shows intelligence discovering the minimum forms of continuity that remain available after disaster.

Survival level	Meaning
Extinction avoidance	Something remains alive. This is the lowest and most basic survival threshold.
Reproductive continuity	The surviving population is not only a relic; births can still occur after collapse.
Refugia stability	Survival exists in regions that retain enough energy, habitability, and local organization to persist.
Distributed survival	A lineage survives in more than one pocket, reducing the chance that one local failure ends it.
Recoverable complexity	The survivor retains enough complexity, memory, or institutions to rebuild higher-order society.
Civilizational continuity	The peak civilization's cooperative, civic, technological, and institutional life remains meaningfully continuous.

Across the study, the first four levels are often achievable. The last two are far less clearly achieved. This is why the final interpretation must not collapse all survival into success.

6. What the Crucible Does Not Show

A closing synthesis also needs restraint. The Crucible is a model world, not an oracle. It shows what happened under a specific engine, specific inherited parameters, specific event models, and specific project conditions. Its conclusions are strongest when stated as model findings and weakest when treated as direct claims about real societies.

Not shown	Why the limit matters
It does not prove inevitable human extinction.	Many scenarios preserve survivors, and some long-horizon cases show low-energy equilibrium.
It does not prove civilizational recovery.	Final states are usually refugia rather than restored high-complexity societies.
It does not prove that any real-world moral trait is good or bad.	Traits such as xeno trust, opportunism, hardiness, and legacy are model variables, not ethical categories.
It does not prove that geography is irrelevant.	Geography changes paths, winners, and refugia location, even when it does not preserve pluralism.
It does not prove climate is the only cause of collapse.	Climate acts as a severity modifier and physical tail layered onto overshoot, competition, and systemic stress.
It does not prove that LLMs save the world.	LLM redesign improves some survival traits and outcompetes controls in some contexts, but mainly by finding collapse-survival architectures.

The appropriate claim is more careful and more useful: inside this world, when agents are allowed to build, compete, evolve, and face delayed consequences, the most stable repeated outcome is not flourishing technological pluralism. It is a movement from emergence to overshoot, from overshoot to collapse, and from collapse to selected refugia.

7. Methodological Meaning

The Crucible's value is not that it supplies a final answer to history. Its value is that it creates a disciplined experimental space where assumptions can be made visible. When geography is changed, the route

changes. When competition is isolated, opponent effects become visible. When designers receive evidence, adaptation becomes visible. When climate reservoirs are strengthened, the future narrows.

The study also gives LLMs an unusual role. They are not merely commentators on the data. In Project 3 and the later design sequence, they become designers inside the experiment: they read evidence, alter founding lineages, and then watch those lineages live or die under the same world engine. This makes The Crucible partly a simulation of agents and partly a simulation of evidence-driven design itself.

That is why the results matter even when they are model-specific. The study does not need to be a prediction engine to be instructive. It shows how difficult it is for intelligence to reason its way out of a world whose physical, social, and reproductive constraints have already been set in motion.

8. Final Closing Reflection

The Crucible begins with a generous possibility: perhaps designed intelligence, given enough information, can create better survival strategies. Across the projects, that possibility is not dismissed. The designed lineages do learn. Some survive where others fail. Some become more coherent, more selective, more durable, and more capable of finding the narrow ledges of post-collapse life.

But the study also makes clear that survival is a smaller victory than it first appears. The most successful lineages often win by accepting a reduced future. They become less expansive, less trusting of outsiders, less reproductive, more internally cohesive, and more suited to endurance than renewal. They are not the civilizations that imagined the peak would last. They are what remains after the peak has ended.

That may be the most important closing insight. The Crucible does not say that collapse is the only possible future, or that intelligence is useless, or that life cannot persist. It says that the form of survival is shaped by the crisis that selects it. If the world rewards hard remnant continuity, then the survivors become hard remnants. If the physical tail is forgiving, equilibrium may return. If the physical tail is brutal, the future narrows to a few pockets, a few lineages, and a few traits strong enough to endure.

The study therefore ends neither in triumph nor despair. It ends with a distinction. To survive is not necessarily to recover. To adapt is not necessarily to become wiser. To build civilization is not necessarily to preserve it. The Crucible shows a world where life can continue after civilization breaks, but where the price of continuation is often simplification.

Appendices

Technical and source appendices moved from the project sections

The following appendices were moved from the individual project sections to the end of the full report so the main narrative reads as one continuous document.

Appendix from Introduction to the Study

Appendix - Engine Detail at a Glance

This appendix provides a compact technical bridge for readers who want more detail before entering the project reports. It is intentionally short; the project documents and query workbooks provide the detailed analysis.

Feature	Concise description
Genome size	33 core inherited numeric founder parameters in the current controller: 3 physical/reproductive traits, 27 social decision values, and 3 wander values.
Social matrix	Trust, forgiveness, and opportunism each contain nine values: K/S/X relation classes crossed with abundance/scarcity/desperation states.
State mapping	Abundance, scarcity, and desperation are determined dynamically from local energy return and agent energy.
Civic identity	Shared residence, institutions, and stability can make co-residents behave more like kin, while stress can weaken civic belonging.
Interaction rounds	Each match uses three Prisoner's Dilemma rounds. The first round is trust-led; later rounds depend on previous partner behaviour.
Regime ascent	Regions advance when complexity exceeds thresholds. Regimes move from Foraging to Agrarian, Industrial, and Technological.
Reproduction	Birth requires maturity, cooldown, energy, fertility probability, and a successful birth attempt. Children inherit with mutation/drift.
Mortality	Agents can die through old age, hazards, starvation, failed early-life conditions, and event-driven pressures.
Climate/event layer	Historical events and addendum modes can alter heat, damage, chaos, hazards, fertility, capacity, regrowth, and energy return.

Appendix from Project 1: Baseline Crucible

Appendix - Analysis Source Workbooks

The analysis content is derived from four Project 1 query-result workbooks. Each workbook contains the tabulated query outputs used to construct the corresponding narrative, charts, tables, and synthesis. Earlier introductory drafts and operational execution notes are treated as archival project materials rather than primary analysis sources.

Workbook	Analysis section	Content used
Project1_SetA_Query_Results.xlsx	Set A - Baseline World	Baseline population, complexity, monopoly, death causes, Prisoner's Dilemma behavior, climate stress, trait signals, and late refugia.
Project1_SetB_Query_Results.xlsx	Set B - Bottleneck World	One-pass geography, winner diversity, transition dynamics, Prisoner's Dilemma behavior, death causes, climate filtering, and final refugia.
Project1_SetC_Query_Results.xlsx	Set C - Duel World	Silent Sting / Gradient Covenant duel outcomes, comparative traits, collapse timing, mortality, social behavior, and final replacement geography.
Project1_SetADE_Query_Results.xlsx	Sets A-D-E climate comparison	Default, harsher, and gentler climate-chaos comparison for population scale, collapse timing, complexity loss, mortality, resilience, tribe sensitivity, and refugia.

Appendix from Closing Synthesis

Appendix - Source Basis

This closing synthesis is based on the five compressed project-note documents prepared from the full project reports, together with the study opening frame. The purpose is to integrate the project arcs at a higher level rather than repeat the detailed conclusions from each individual report.

Source note	Used for
Project 1 Baseline Crucible Notes	Baseline world, bottleneck geography, duel world, climate-severity comparison, and Project 1 synthesis.

Project 2 Competition Notes	Solo viability, all-LLM duels, triad structure, competition dynamics, genetics, and mono-tribal refugia.
Project 3 Evolution Notes	Evidence-driven redesign, LLM generations, trait convergence, biologic controls, and adaptation limits.
Project 4 Bombs Notes	Aerosol demasking, methane release, nuclear meltdown, stacked consequences, and survival-tier distinction.
Epilogue Notes	Obsidian reference, Ares design, network collapse, long-horizon survival, reservoir calibration, and final reservoir projection.

Appendix from Epilogue

Phase 1 examines why the Obsidian Phalanx Refuge G3 becomes a recurring survival reference in the late Crucible phases. The key interpretive theory is that Obsidian changes the survival problem from maintaining civilization through collapse to producing enough hardened, low-metabolism survivors that collapse leaves viable refugia behind.

Query appendix

This appendix lists the workbook tabs that support the phase analysis. The tab names correspond to query outputs, and the purpose column describes the role each output plays in the evidence trail.

Workbook tab	Purpose
00_obsidian_identity_and_lineag	Identifies Obsidian's lineage, player identifiers, and design context.
01_obsidian_cross_phase_record	Collects Obsidian's epoch-level performance across the stress phases.
02_cross_phase_survival_ranking	Ranks all G3 competitors by survival rate and survival scale.
03_phase_by_phase_competitor_co	Compares Obsidian, other language-model tribes, and biologic tribes by phase.
04_founder_trait_distinctivenes	Compares founder traits such as hardiness, vitality, legacy, trust, opportunism, and wander.
05_founder_to_elite_drift	Measures how founder traits drift into elite survivor traits.
06_obsidian_timeline_by_phase	Tracks Obsidian's timeline by phase and epoch.
07_obsidian_collapse_window_con	Examines Obsidian during the crash window around peak and collapse.
08_final_refugia_obsidian_and_w	Maps final refugia and dominant final regions.
09_late_reproductive_viability	Tests whether late survivors continue to reproduce after the crash.
10_event_audit_for_obsidian_pha	Audits historical event conditions used in the phase set.

11_objective_notes_for_all_g3_l	Preserves designer notes and objectives for G3 language-model tribes.
12_Ares_design_lesson_scorecar	Summarizes design lessons used to motivate Ares.

Phase 3 functions as the first Epilogue field test after the Ares design intervention.

Query appendix

This appendix lists the workbook tabs that support the phase analysis. The tab names correspond to query outputs, and the purpose column describes the role each output plays in the evidence trail.

Workbook tab	Purpose
00_setup_persistent_helper_tabl	Builds helper records for the Ares field-test analysis.
01_full_series_scoreboard	Ranks all tribes in epochs 181-190 by survival, final population, peak population, energy, and regime reach.
02_Ares_epoch_by_epoch_with_wi	Shows Ares by epoch, including wins, losses, peak scale, final population, and winner identity.
03_Ares_loss_epochs_only	Isolates Ares loss epochs for failure diagnosis.
04_Ares_wins_vs_losses_summary	Compares Ares wins against Ares losses across demographic, energy, and mortality measures.
05_Ares_vs_obsidian_by_epoch	Compares Ares and Obsidian directly in each epoch.
06_who_beat_Ares_in_losses	Identifies which competitors survived or defeated Ares in the loss cases.
07_Ares_loss_population_trajec	Tracks Ares population and energy trajectories in its loss epochs.
08_Ares_phase_window_trajector	Summarizes Ares through key pulse windows.
09_Ares_losses_system_conditio	Examines world conditions in Ares loss epochs.
10_wins_vs_losses_system_condit	Compares world conditions in Ares wins and losses.
11_Ares_regime_gate_check	Tests whether Ares failures came from regime-gate failure.
12_Ares_elite_traits_wins_vs_l	Compares elite traits in Ares wins and losses.
13_late_refugia_dominant_tribes	Shows late refugia dominance and final region ownership.
14_late_internet_contact_persis	Checks late contact and network persistence during the refugia period.
15_compact_failure_explanation_	Explores why compact/refugia designs failed or survived relative to Ares.

Phase 4 clarifies whether post-collapse plural survival is prevented by late network contact or by earlier sorting dynamics.

Query appendix

This appendix lists the workbook tabs that support the phase analysis. The tab names correspond to query outputs, and the purpose column describes the role each output plays in the evidence trail.

Workbook tab	Purpose
00_setup_persistent_helper_tabl	Builds helper records for network-collapse analysis.
01_active_field_and_event_audit	Audits active tribes, events, and run conditions for epochs 191-200.
02_survivor_count_by_epoch	Counts final surviving tribes by epoch.
03_series_scoreboard_191_200	Ranks tribe performance across the natural network-collapse series.
04_final_survivors_detail	Lists final survivors, final population, peak population, and regime reach.
05_baseline_vs_network_collapse	Compares Phase 4 against the Phase 3 baseline.
06_late_network_contact_proxy_b	Measures late-stage contact proxies by window and series.
07_contact_proxy_baseline_vs_ne	Compares contact proxies between Phase 3 and Phase 4 conditions.
08_epoch_200_two_survivor_diagn	Diagnoses the key two-survivor epoch.
09_epoch_200_survivor_trajector	Tracks trajectory of each epoch 200 final survivor.
10_alive_tribe_count_timeline	Tracks living-tribe counts through time.
11_first_single_or_multi_surviv	Finds when each epoch first reaches single-survivor or multi-survivor status.
12_final_refugia_region_map	Maps final populated refugia regions.
13_final_refugia_summary_by_epo	Summarizes final refugia by epoch.
14_epoch_200_final_refugia_spat	Shows spatial separation and dominance in epoch 200 final refugia.
15_world_population_slope_40000	Checks late population slope near the endpoint.
16_late_birth_death_balance_sur	Measures late birth/death balance for survivors.
17_winner_predictability_and_su	Tests whether final winners were predictable from earlier scale or rank.

Phase 5 checks whether the older endpoint was stopping in the middle of extinction or capturing the beginning of a stable low-energy future.

Query appendix

This appendix lists the workbook tabs that support the phase analysis. The tab names correspond to query outputs, and the purpose column describes the role each output plays in the evidence trail.

Workbook tab	Purpose
00_setup_persistent_helper_tabl	Builds helper records for 80,000-pulse extension analysis.
01_active_field_and_event_audit	Audits active field, events, and conditions for epochs 201-210.
02_final_survivor_count_80k	Counts final survivors at the 80,000-pulse endpoint.
03_series_scoreboard_201_210	Ranks tribe performance across the 80,000-pulse series.
04_40k_snapshot_vs_80k_final_by	Compares the 40,000-pulse state with the 80,000-pulse endpoint.
05_alive_tribe_count_timeline_8	Tracks living-tribe counts through the extended run.
06_first_equilibrium_pulse	Finds when the final survivor count is reached.
07_final_survivors_detail_80k	Details final survivors at 80,000 pulses.
08_winner_variability_191_200_v	Compares winner distribution to Phase 4.
09_post40k_birth_death_balance	Measures post-40,000 demographic balance by window.
10_climate_by_epoch_and_window	Summarizes climate variables by epoch and time window.
11_climate_40k_to_80k_delta	Measures climate changes between 40,000 and 80,000 pulses.
12_climate_after_population_col	Tests climate behavior after population collapse.
13_emissions_and_advanced_pop_a	Tracks emissions and advanced population after collapse.
14_network_contact_after_40k	Checks whether network/contact interactions persist after 40,000 pulses.
15_final_refugia_region_map_80k	Maps final populated refugia at 80,000 pulses.
16_final_refugia_summary_80k	Summarizes final refugia at 80,000 pulses.
17_winner_path_40k_to_80k	Tracks each winner's path from 40,000 to 80,000 pulses.
18_predictability_of_80k_winner	Tests whether 80,000-pulse winners were predictable from earlier state.
19_climate_tuning_signal_summar	Summarizes climate-tail signals that motivated reservoir addendum work.

Phase 7 produces the intended experimental shape: every epoch reaches a large civilization-scale boom, reaches Regime 4 (Technological) at least briefly, arms the reservoirs at pulse 31,000, crashes almost immediately after the peak, and ends at pulse 80,000 in tiny Regime 2 (Agrarian) refugia.

Query appendix

This appendix lists the workbook tabs that support the phase analysis. The tab names correspond to query outputs, and the purpose column describes the role each output plays in the evidence trail.

Workbook tab	Purpose
00_series_metadata_check	Confirms engine version, mode, network-collapse flag, and archived metadata.
01_epoch_outcomes_and_climate_s	Summarizes boom, crash, reservoir arming, and final state by epoch.
02_collapse_milestones_after_pe	Measures how quickly each epoch collapses after peak population.
03_reservoir_timing_and_compone	Verifies reservoir timing and component behavior.
04_climate_pressure_windows	Summarizes climate and demographic pressure by time window.
05_survivors_by_epoch	Counts survival and winning frequency by tribe.
06_survivor_frequency_by_tribe	Measures how many tribes survive in each epoch.
07_survivor_counts_by_epoch	Scores tribe life histories across births, deaths, energy, and regime reach.
08_tribe_life_history_scoreboar	Tests whether final winners were predictable from boom rankings.
09_winner_predictability_rankin	Compares founder genetics for survivors and extinct tribes.
10_founder_genetics_survivors_v	Tests relative trait thresholds such as high hardiness, high vitality, and low legacy.
11_founder_trait_thresholds	Compares elite/representative genetics by final role.
12_elite_genetics_by_final_role	Shows tribe state at key pulses before, during, and after collapse.
13_tribe_population_at_key_puls	Aggregates post-crash survival metrics by tribe and window.
14_postcrash_survival_windows_b	Measures identity, contact, and Prisoner's Dilemma interaction patterns by window.
15_global_interaction_windows	Lists final populated regions and dominant tribes.
16_final_refugia_by_region	Summarizes final refugia by epoch.
17_refugia_summary_by_epoch	Shows which tribes dominate final refugia regions.
18_region_dominance_by_tribe	Compares language-model and biologic tribe outcomes.
19_llm_vs_biologic_outcomes	Tracks winners and outcome markers by epoch.
20_epoch_to_epoch_winner_trends	Lists strategy notes for tribes that survived at least once.
21_strategy_notes_for_survivors	Compares founder traits for winners, other survivors, and extinct tribes.
22_survival_vs_founder_traits_s	Measures post-peak collapse milestones.

Addendum: Lessons for a Future Crucible Study

1. Geography mode should be explicitly versioned by project

Project 3 introduced quadrant-barrier topology. Later projects should either have reverted to the baseline mountain/lake geography or explicitly declared the continued use of Project 3 topology. Future studies should archive a per-region terrain mask so rendered maps and engine geography can always be audited.

2. Terrain should be archived as data, not only generated in code

Mountains, lakes, barriers, fertile regions, and passable edges should be stored per epoch. That would make later visualization and analysis cleaner.

3. Founder uniqueness should not be forced

The original LLM design prompt asked for unique agents. In a future study, models should be allowed to duplicate founder profiles if they believe repetition, standardization, or cloning is strategically useful.

4. Prompt freedoms should be stated more explicitly

The models should be told whether they are optimizing for survival, dominance, cooperation, resilience, civilization, or any self-chosen goal. If the goal is open-ended, that should be explicit in the prompt archive.

5. Experimental switches should have project-level reset checks

After a special project changes geography, climate, network behaviour, or agent logic, the next project should begin with a checklist confirming which switches remain active.