The Crucible

Project 2 - Competition

Research Plan, Execution Design, and Outcome Analysis

One-on-one, triad, and single-tribe experiments using a fresh all-LLM tribe roster

June 17, 2026

Contents

Part I - Research Plan and Execution Design

Project 2 is the Crucible's competition project. It moves from the baseline world behavior 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 Fractal Phoenix	99, 97	5	Test the second edge of the A/B/C 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 refugia	Industrial civilization occurred, collapsed, and simplified rather than 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 summaries, 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.

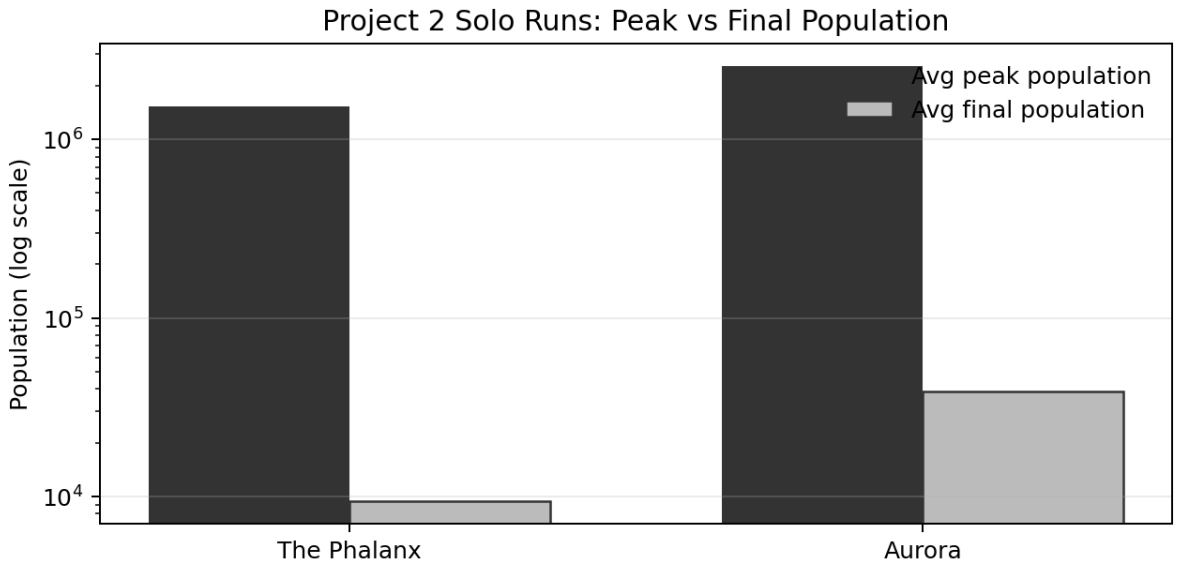
1. Data Summary and Workbook Coverage

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.



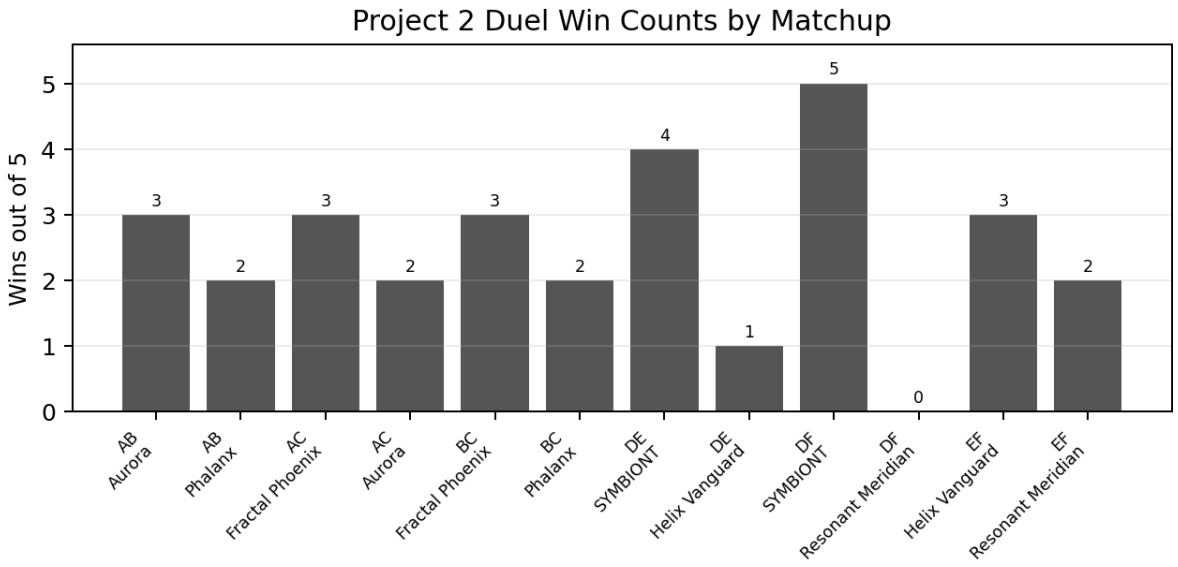
The chart compares average peak and final population for the two solo blocks on a logarithmic scale. Aurora reaches a much larger average peak and final population, while The Phalanx survives but leaves a thinner refugia system.

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.



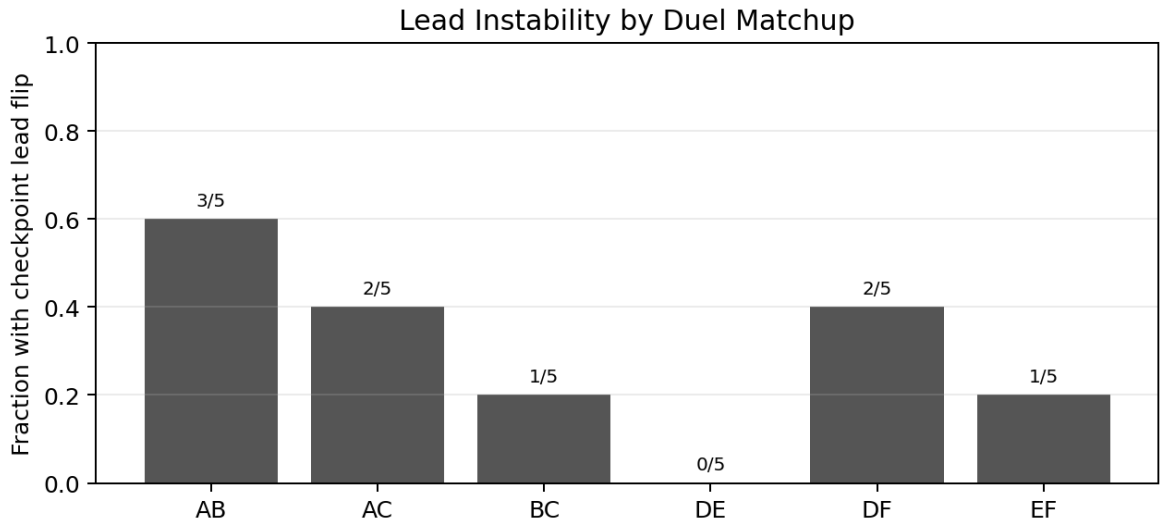
The chart shows the number of wins on each five-epoch duel edge. Every first-triad matchup is split 3-2. The clearest dominance appears in SYMBIONT-ZERO vs The Resonant Meridian, where SYMBIONT-ZERO wins 5-0.

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.

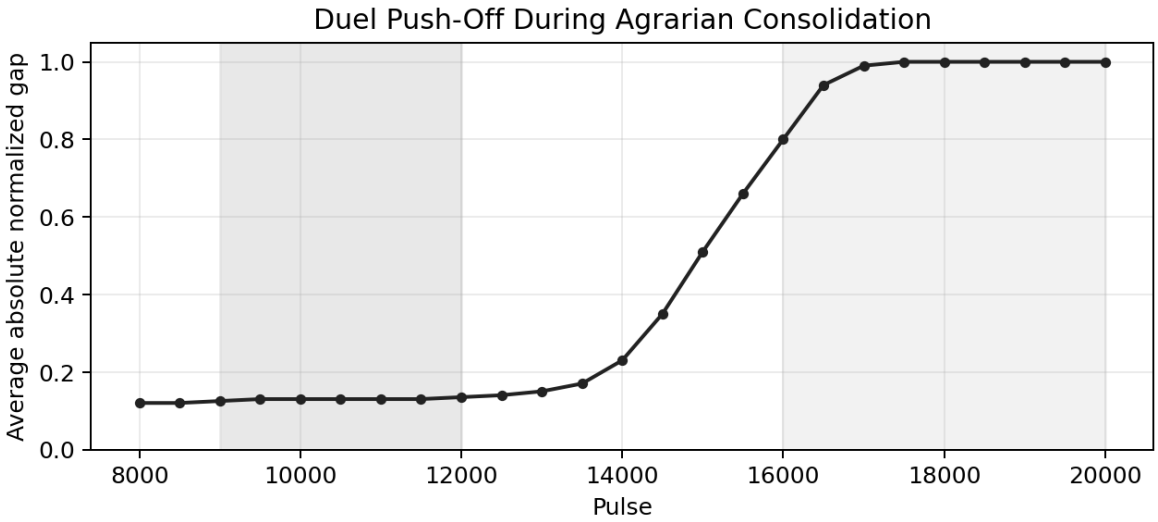
4. Competition Dynamics: Push-Off and Rank Reversal

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%

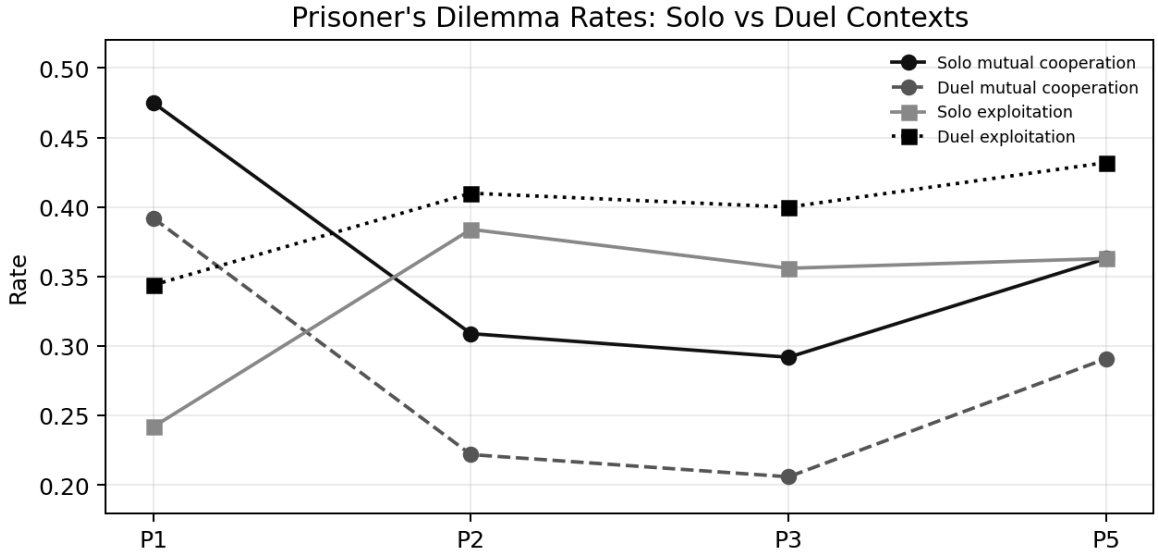


The line shows the average absolute normalized population gap across duel epochs from pulse 8,000 to 20,000. The gap is modest early, then widens sharply during the push-off and consolidation phase.

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.



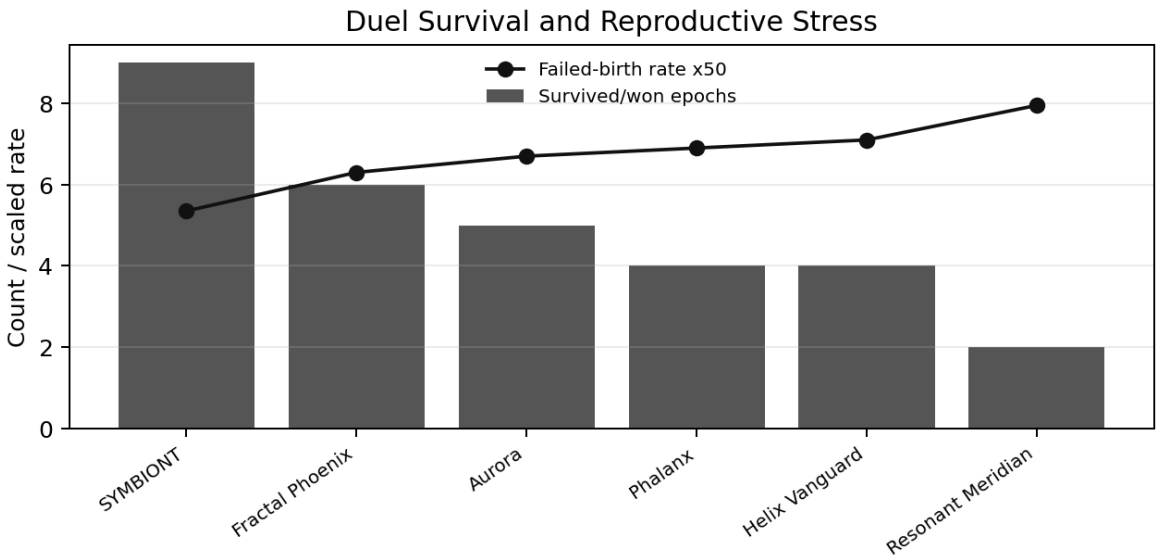
Duels are less cooperative and more exploitative than solo worlds during the push-off/consolidation phase. The social world hardens when the losing tribe is most vulnerable.

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.



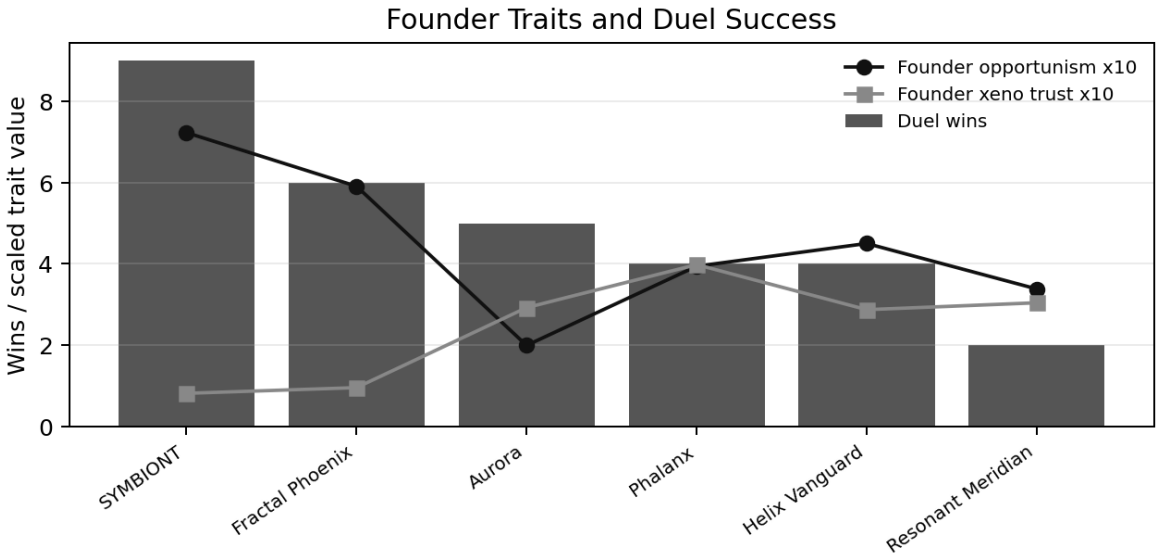
SYMBIONT-ZERO combines the strongest duel performance with one of the lowest failed-birth rates, while weaker duel performers show higher reproductive stress.

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 Resonant Meridian	29.9	31.2	38.9	0.304	0.338	0.275	2	6,995

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

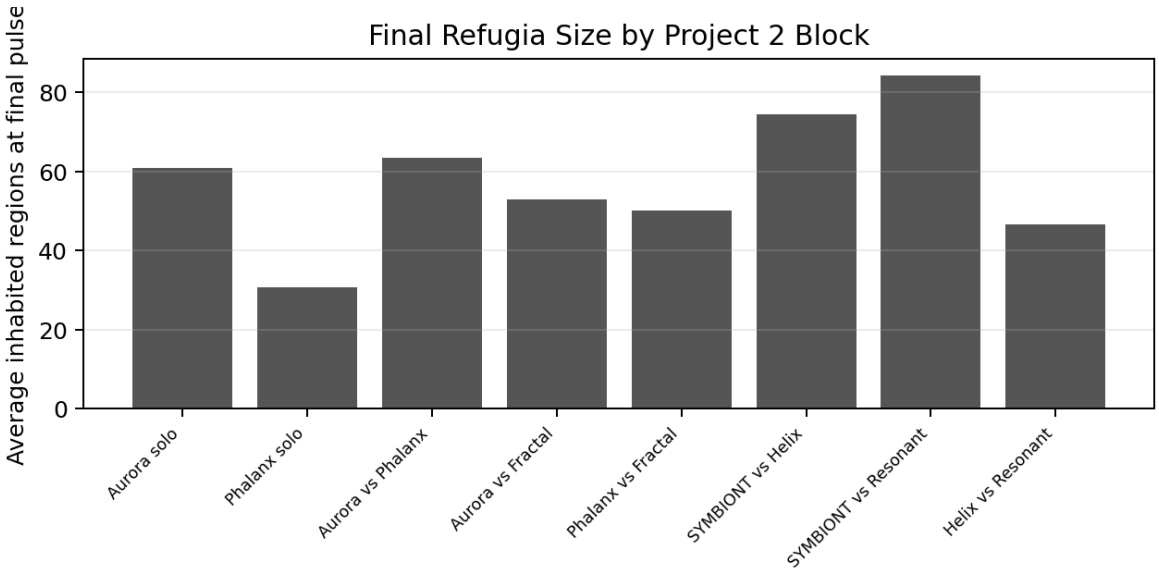
The elite-genetics sheet contains 1,000 records sampled from top survivors. Across all six tribes, elites have substantially higher hardiness than founders. This appears to be one of the most consistent internal evolutionary signals in the Project 2 data. Xeno trust generally declines among elites for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