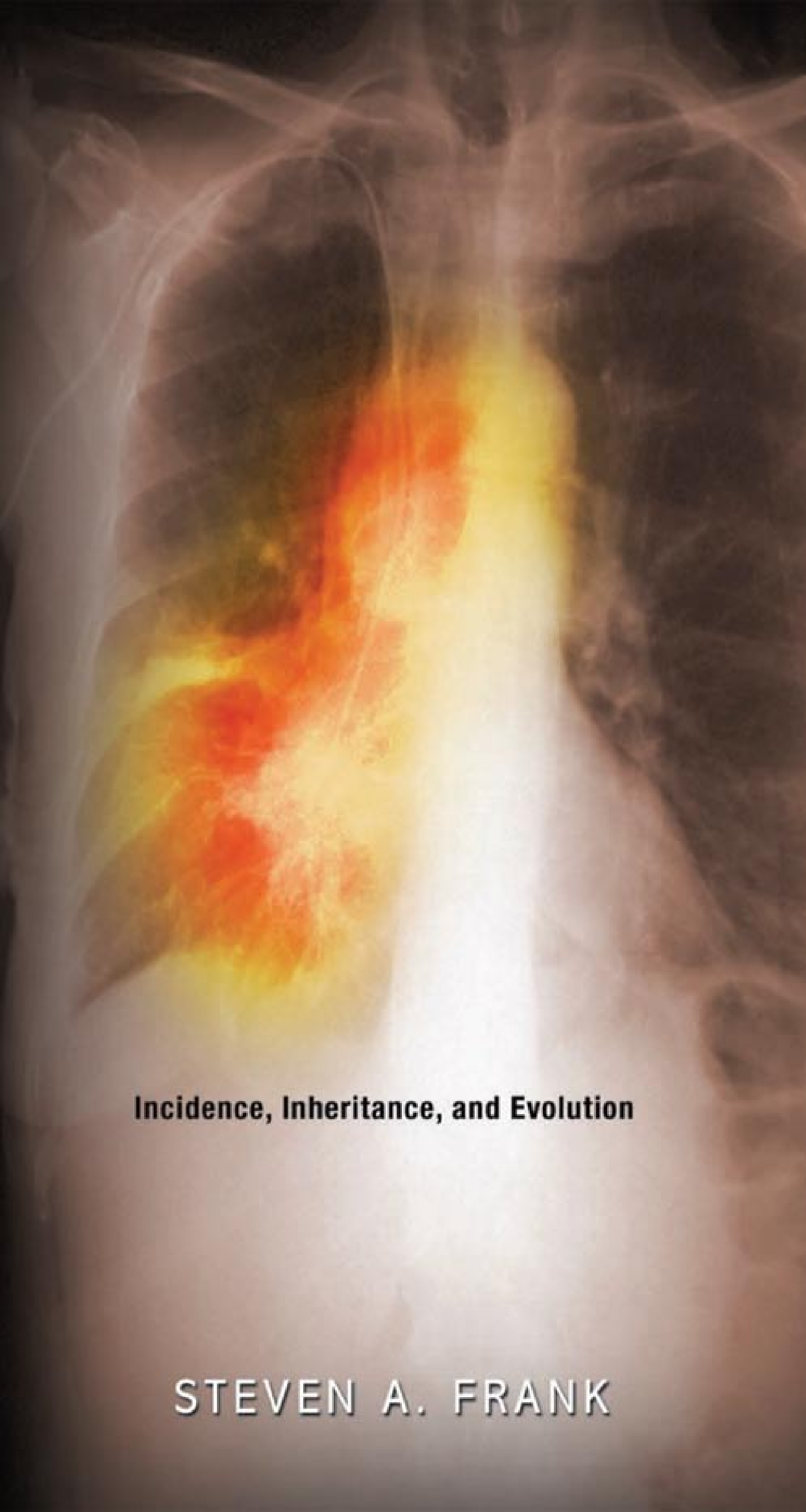
DYNAMICS OF CANCER



Incidence, Inheritance, and Evolution

STEVEN A. FRANK

Dynamics of Cancer

Incidence, Inheritance, and Evolution

STEVEN A. FRANK

Princeton University Press
Princeton and Oxford

Copyright © 2007 by Steven A. Frank
Published by Princeton University Press
41 William Street, Princeton, New Jersey 08540
In the United Kingdom: Princeton University Press
3 Market Place, Woodstock, Oxfordshire OX20 1SY

This work is provided under the terms of this Creative Commons Public License (CCPL). You may use the work in accordance with the terms of the CCPL. This work is otherwise protected by copyright and/or other applicable law. http://creativecommons.org/licenses/by-nc/2.5/.

Library of Congress Cataloging-in-Publication Data

Frank, Steven A., 1957–
Dynamics of cancer : incidence, inheritance, and evolution / Steven A. Frank. p. ; cm. – (Princeton series in evolutionary biology)
Includes bibliographic references and index.
ISBN 978–0–691–13365–2 (cloth : alk. paper)
ISBN 978–0–691–13366–9 (pbk. : alk. paper)
1. Carcinogenesis. 2. Cancer–Age factors.
3. Cancer–Genetic aspects. 4. Cancer–Epidemiology.
I. Title. II. Series.
[DNLM: 1. Neoplasms–etiology. 2. Age of Onset.
3. Gene Expression Regulation, Neoplastic. 4. Genetic Predisposition to Disease. 5. Mutagenesis. 6. Stem Cells. QZ 202 F828d 2007]
RC268.5.F63 2007
616.99′4071–dc22 2007007855

British Library Cataloging-in-Publication Data is available

Typeset by the author with TEX
Composed in Lucida Bright

Printed on acid-free paper. ∞
press.princeton.edu
Printed in the United States of America
10 9 8 7 6 5 4 3 2 1

I wish I had the voice of Homer
To sing of rectal carcinoma,
Which kills a lot more chaps, in fact,
Than were bumped off when Troy was sacked.
—J. B. S. Haldane

Contents

1 Introduction 1

1.1 Aims 2
1.2 How to Read 4
1.3 Chapter Summaries 5

PART I: BACKGROUND

2 Age of Cancer Incidence 17

2.1 Incidence and Acceleration 19
2.2 Different Cancers 20
2.3 Childhood Cancers 23
2.4 Inheritance 25
2.5 Carcinogens 29
2.6 Sex Differences 32
2.7 Summary 35

3 Multistage Progression 36

3.1 Terminology 37
3.2 What Is Multistage Progression? 38
3.3 Multistage Progression in Colorectal Cancer 39
3.4 Alternative Pathways to Colorectal Cancer 43
3.5 Changes during Progression 49
3.6 What Physical Changes Drive Progression? 50
3.7 What Processes Change during Progression? 51
3.8 How Do Changes Accumulate in Cell Lineages? 55
3.9 Summary 58

4 History of Theories 59
4.1 Origins of Multistage Theory 61
4.2 A Way to Test Multistage Models 65
4.3 Cancer Is a Genetic Disease 69
4.4 Can Normal Somatic Mutation Rates Explain Multistage Progression? 71
4.5 Clonal Expansion of Premalignant Stages 74
4.6 The Geometry of Cell Lineages 76
4.7 Hypermutation, Chromosomal Instability, and Selection 78
4.8 Epigenetics: Methylation and Acetylation 79
4.9 Summary 80

PART II: DYNAMICS

5 Progression Dynamics 85
5.1 Background 86
5.2 Observations to Be Explained 89
5.3 Progression Dynamics through Multiple Stages 90
5.4 Why Study Quantitative Theories? 93
5.5 The Basic Model 93
5.6 Technical Definitions of Incidence and Acceleration 94
5.7 Summary 95
6 Theory I 96
6.1 Approach 97
6.2 Solution with Equal Transition Rates 97
6.3 Parallel Evolution within Each Individual 100
6.4 Unequal Transition Rates 103

6.5 Time-Varying Transition Rates 109
6.6 Summary 114

7 Theory II 115

7.1 Multiple Pathways of Progression 116
7.2 Discrete Genetic Heterogeneity 120
7.3 Continuous Genetic and Environmental Heterogeneity 129
7.4 Weibull and Gompertz Models 136
7.5 Weibull Analysis of Carcinogen Dose-Response Curves 139
7.6 Summary 142

8 Genetics of Progression 143

8.1 Comparison between Genotypes in Human Populations 144
8.2 Comparison between Genotypes in Laboratory Populations 154
8.3 Polygenic Heterogeneity 160
8.4 Summary 164

9 Carcinogens 165

9.1 Carcinogen Dose-Response 166
9.2 Cessation of Carcinogen Exposure 180
9.3 Mechanistic Hypotheses and Comparative Tests 190
9.4 Summary 201

10 Aging 202

10.1 Leading Causes of Death 203
10.2 Multistage Hypotheses 206
10.3 Reliability Models 207
10.4 Conclusions 209
10.5 Summary 209

PART III: EVOLUTION

11 Inheritance 213

11.1 Genetic Variants Affect Progression and Incidence 214

11.2 Progression and Incidence Affect Genetic Variation 234

11.3 Few Common or Many Rare Variants? 243

11.4 Summary 250

12 Stem Cells: Tissue Renewal 251

12.1 Background 252

12.2 Stem-Transit Program of Renewal 253

12.3 Symmetric versus Asymmetric Stem Cell Divisions 264

12.4 Asymmetric Mitoses and the Stem Line Mutation Rate 265

12.5 Tissue Compartments and Repression of Competition 269

12.6 Summary 270

13 Stem Cells: Population Genetics 271

13.1 Mutations during Development 272

13.2 Stem-Transit Design 280

13.3 Symmetric versus Asymmetric Mitoses 283

13.4 Summary 285

14 Cell Lineage History 286

14.1 Reconstructing Cellular Phylogeny 287

14.2 Demography of Progression 295

14.3 Somatic Mosaicism 304

14.4 Summary 308

15 Conclusions 309

Appendix: Incidence 314

References 335

Author Index 361

Subject Index 373

1

Introduction

Through failure we understand biological design. Geneticists discover the role of a gene by studying how a mutation causes a system to fail. Neuroscientists discover mental modules for face recognition or language by observing how particular brain lesions cause cognitive failure.

Cancer is the failure of controls over cellular birth and death. Through cancer, we discover the design of cellular controls that protect against tumors and the architecture of tissue restraints that slow the progress of disease.

Given a particular set of genes and a particular environment, one cannot say that cancer will develop at a certain age. Rather, failure happens at different rates at different ages, according to the age-specific incidence curve that defines failure.

To understand cancer means to understand the genetic and environmental factors that determine the incidence curve. To learn about cancer, we study how genetic and environmental changes shift the incidence curve toward earlier or later ages.

The study of incidence means the study of rates. How does a molecular change alter the rate at which individuals progress to cancer? How does an inherited genetic change alter the rate of progression? How does natural selection shape the design of regulatory processes that govern rates of failure?

Over fifty years ago, Armitage and Doll (1954) developed a multistage theory to analyze rates of cancer progression. That abstract theory turned on only one issue: ultimate system failure—cancer—develops through a sequence of component failures. Each component failure, such as loss of control over cellular death or abrogation of a critical DNA repair pathway, moves the system one stage along the progression to disease. Rates of component failure and the number of stages in progression determine the age-specific incidence curve. Mutations that knock out a component or increase the rate of transition between stages shift the incidence curve to earlier ages.

I will review much evidence that supports the multistage theory of cancer progression. Yet that support often remains at a rather vague level: little more than the fact that progression seems to follow through multiple stages. A divide separates multistage theory from the daily work of cancer research.

The distance between theory and ongoing research arose naturally. The theory follows from rates of component failures and age-specific incidence in populations; most cancer research focuses on the mechanistic and biochemical controls of particular components such as the cell cycle, cell death, DNA repair, or nutrient acquisition. It is not easy to tie failure of a particular pathway in cell death to an abstract notion of the rate of component failure and advancement by a stage in cancer progression.

In this book, I work toward connecting the great recent progress in molecular and cellular biology to the bigger problem: how failures in molecular and cellular components determine rates of progression and the age-specific incidence of cancer. I also consider how one can use observed shifts in age-specific incidence to analyze the importance of particular molecular and cellular aberrations. Shifts in incidence curves measure changes in failure rates; changes in failure rates provide a window onto the design of molecular and cellular control systems.

1.1 Aims

The age-specific incidence curve reflects the processes that drive disease progression, the inheritance of predisposing genetic variants, and the consequences of carcinogenic exposures. It is easy to see that these various factors must affect incidence. But it is not so obvious how these factors alter measurable, quantitative properties of age-specific incidence.

My first aim is to explore, in theory, how particular processes cause quantitative shifts in age-specific incidence. That theory provides the tools to develop the second aim: how one can use observed changes in age-specific incidence to reveal the molecular, cellular, inherited, and environmental factors that cause disease. Along the way, I will present a comprehensive summary of observed incidence patterns, and I will synthesize the intellectual history of the subject.

I did not arbitrarily choose to study patterns of age-specific incidence. Rather, as I developed my interests in cancer and other age-related diseases, I came to understand that age-specific incidence forms the nexus through which hidden process flows to observable outcome. In this book, I address the following kinds of questions, which illustrate the link between disease processes and age-related outcomes.

Faulty DNA repair accelerates disease onset—that is easy enough to guess—but does poor repair accelerate disease a little or a lot, early in life or late in life, in some tissues but not in others?

Carcinogenic chemicals shift incidence to earlier ages: one may reasonably measure whether a particular dosage is carcinogenic by whether it causes a shift in age-specific incidence, and measure potency by the degree of shift in the age-incidence curve. Why do some carcinogens cause a greater increase in disease if applied early in life, whereas other carcinogens cause a greater increase if applied late in life? Why do many cancers accelerate rapidly with increasing time of carcinogenic exposure, but accelerate more slowly with increasing dosage of exposure? What processes of disease progression do the chemicals affect, and how do changes in those biochemical aspects of cells and tissues translate into disease progression?

Inherited mutations sometimes abrogate key processes of cell cycle control or DNA repair, leading to a strong predisposition for cancer. Why do such mutations shift incidence to earlier ages, but reduce the rate at which cancer increases (accelerates) with age?

Why do the incidences of most diseases, including cancer, accelerate more slowly later in life? What cellular, physiological, and genetic processes of disease progression inevitably cause the curves of death to flatten in old age?

Inherited mutations shift incidence to earlier ages. How do the particular changes in age-specific incidence caused by a mutation affect the frequency of that mutation in the population?

How do patterns of cell division, tissue organization, and tissue renewal via stem cells affect the accumulation of somatic mutations in cell lineages? How do the rates of cell lineage evolution affect disease progression? How do alternative types of heritable cellular changes, such as DNA methylation and histone modification, affect progression? How can one measure cell lineage evolution within individuals?

I will not answer all of these questions, but I will provide a comprehensive framework within which to study these problems.

Above all, this book is about biological reliability and biological failure. I present a full, largely novel development of reliability theory that accounts for biological properties of variability, inheritance, and multiple pathways of disease. I discuss the consequences of reliability and failure rates for evolutionary aspects of organismal design. Cancer provides an ideal subject for the study of reliability and failure, and through the quantitative study of failure curves, one gains much insight into cancer progression and the ways in which to develop further studies of cancer biology.

1.2 How to Read

Biological analysis coupled with mathematical development can produce great intellectual synergy. But for many readers, the mixed language of a biology-math marriage can seem to be a private dialect understood by only a few intimates.

Perhaps this book would have been an easier read if I had published the quantitative theory separately in journals, and only summarized the main findings here in relation to specific biological problems. But the real advance derives from the interdisciplinary synergism, diluted neither on the biological nor on the mathematical side. If fewer can immediately grasp the whole, more should be attracted to try, and with greater ultimate reward. Progress will ultimately depend on advances in biology, on advances in the conceptual understanding of reliability and failure, and on advances in the quantitative analysis and interpretation of data.

I have designed this book to make the material accessible to readers with different training and different goals. Chapters 2 and 3 provide background on cancer that should be accessible to all readers. Chapter 4 presents a novel historical analysis of the quantitative study of age-specific cancer incidence. Chapter 5 gives a gentle introduction to the quantitative theory, why such theory is needed, and how to use it. That mathematical introduction should be readable by all.

Chapters 6 and 7 develop the mathematical theory, with much original work on the fundamental properties of reliability and failure in biological systems. Each section in those two mathematical chapters includes a nontechnical introduction and conclusion, along with figures that illustrate the main concepts. Those with allergy to mathematics can glance briefly at the section introductions, and then move along quickly before the reaction grows too severe. The rest of the book applies the quantitative concepts of the mathematical chapters, but does so in a way that can be read with nearly full understanding independently of the mathematical details.

Chapters 8, 9, and 10 apply the quantitative theory to observed patterns of age-specific incidence. I first test hypotheses about how inherited, predisposing genotypes shift the age-specific incidence of cancer. I then evaluate alternative explanations for the patterns of age-specific cancer onset in response to chemical carcinogen exposure. Finally, I analyze data on the age-specific incidence of the leading causes of death, such as heart disease, cancer, cerebrovascular disease, and so on.

I then turn to various evolutionary problems. In Chapter 11, I evaluate the population processes by which inherited genetic variants accumulate and affect predisposition to cancer. Chapters 12 and 13 discuss how somatic genetic mutations arise and affect progression to disease. For somatic cell genetics, the renewal of tissues through tissue-specific adult stem cells plays a key role in defining the pattern of cell lineage history and the accumulation of somatic mutations. Chapter 14 finishes by describing empirical methods to study cell lineages and the accumulation of heritable change.

The following section provides an extended summary of each chapter. I give those summaries so that readers with particular interests can locate the appropriate chapters and sections, and quickly see where I present specific analyses and conclusions. The extended summaries also allow one to develop a customized reading strategy in order to focus on a particular set of topics or approaches. Many readers will prefer to skip the summaries for now and move directly to Chapter 2.

1.3 Chapter Summaries

Part I of the book provides background in three chapters: incidence, progression, and conceptual foundations. Each chapter can be read independently as a self-contained synthesis of a major topic.

Chapter 2 describes the age-specific incidence curve. That failure curve defines the outcome of particular genetic, cellular, and environmental processes that lead to cancer. I advocate the acceleration of cancer as the most informative measure of process: acceleration measures how fast the incidence (failure) rate changes with age. I plot the incidence and acceleration curves for 21 common cancers. I include in the Appendix detailed plots comparing incidence between the 1970s and 1990s, and comparing incidence between the USA, Sweden, England, and Japan. I also compare incidence between males and females for the major cancers.

I continue Chapter 2 with summaries of incidence of major childhood cancers and of inherited cancers. I finish with a description of how chemical carcinogens alter age-specific incidence. Taken together, this chapter provides a comprehensive introduction to the observations of cancer incidence, organized in a comparative way that facilitates analysis of the factors that determine incidence.

Chapter 3 introduces cancer progression as a sequence of failures in components that regulate cells and tissues. I review the different ways in which the concept of multistage progression has been used in cancer research. I settle on *progression* in the general sense of development through multiple stages, with emphasis on how rates of failure for individual stages together determine the observed incidence curve. I then describe multistage progression in colorectal cancer, the clearest example of distinct morphological and genetical stages in tumor development. Interestingly, colorectal cancer appears to have alternative pathways of progression through different morphological and genetic changes; the different pathways are probably governed by different rate processes.

The second part of Chapter 3 focuses on the kinds of physical changes that occur during progression. Such changes include somatic mutation, chromosomal loss and duplication, genomic rearrangements, methylation of DNA, and changes in chromatin structure. Those physical changes alter key processes, resulting, for example, in a reduced tendency for cell suicide (apoptosis), increased somatic mutation and chromosomal instability, abrogation of cell-cycle checkpoints, enhancement of cell-cycle accelerators, acquisition of blood supply into the developing tumor, secretion of proteases to digest barriers against invasion of other tissues, and neglect of normal cellular death signals during migration into a foreign tissue. I finish with a discussion of how changes accumulate over time, with special attention to the role of evolving cell lineages throughout the various stages of tumor development.

Chapter 4 analyzes the history of theories of cancer incidence. I start with the early ideas in the 1920s about multistage progression from chemical carcinogenesis experiments. I follow with the separate line of mathematical multistage theory that developed in the 1950s to explain the patterns of incidence curves. Ashley (1969a) and Knudson (1971) provided the most profound empirical test of multistage progression. They reasoned that if somatic mutation is the normal cause of progression, then individuals who inherit a mutation would have one less step to pass before cancer arises. By the mathematical theory, one less step shifts the incidence curve to earlier ages and reduces the slope (acceleration) of failure. Ashley (1969a) compared incidence in normal individuals and those who inherit a single mutation predisposing to colon cancer: he found the predicted shift in incidence to earlier ages among the predisposed individuals. Knudson (1971) found the same predicted shift between inherited and noninherited cases of retinoblastoma.

I continue Chapter 4 with various developments in the theory of multistage progression. One common argument posits that somatic mutation alone pushes progression too slowly to account for incidence; however, the actual calculations remain ambiguous. Another argument emphasizes the role of clonal expansion, in which a cell at an intermediate stage divides to produce a clonal population that shares the changes suffered by the progenitor cell. The large number of cells in a clonal population raises the target size for the next failure that moves progression to the following stage. I then discuss various consequences of cell lineage history and processes that influence the accumulation of change in lineages. I end by returning to the somatic mutation rate, and how various epigenetic changes such as DNA methylation or histone modification may augment the rate of heritable change in cell lineages.

Part II turns to the dynamics of progression and the causes of the incidence curve. I first present extensive, original developments of multistage theory. I then apply the theory to comparisons between different genotypes that predispose to cancer and to different treatments of chemical carcinogens. I also apply the quantitative theory of age-specific failure to other causes of death besides cancer; the expanded analysis provides a general theory of aging.

Chapter 5 sets the background for the quantitative analysis of incidence. Most previous theory fit specific models to the data of incidence curves. However, fitting models to the data provides almost no insight;

such fitting demonstrates only sufficient mathematical malleability to be shaped to particular observations. A good framework and properly formulated hypotheses express comparative predictions: how incidence shifts in response to changes in genetics and changes in the cellular mechanisms that control rates of progression. This book strongly emphasizes the importance of comparative hypotheses in the analysis of incidence curves and the mechanisms that protect against failure.

I continue Chapter 5 with the observations of incidence to be explained. I follow with simple formulations of theories to introduce the basic approach and to show the value of quantitative theories in the analysis of cancer. I finish with technical definitions of incidence and acceleration, the fundamental measures for rates of failure and how failure changes with age.

Chapters 6 and 7 provide full development of the quantitative theory of incidence curves. Each section begins with a summary that explains in plain language the main conceptual points and conclusions. After that introduction, I provide mathematical development and a visual presentation in graphs of the key predictions from the theory.

In Chapters 6 and 7, I include several original mathematical models of incidence. I developed each new model to evaluate the existing data on cancer incidence and to formulate appropriate hypotheses for future study. These chapters provide a comprehensive theory of age-specific failure, tailored to the problem of multistage progression in cell lineages and in tissues, and accounting for inherited and somatic genetic heterogeneity. I also relate the theory to classical models of aging given by the Gompertz and Weibull formulations. Throughout, I emphasize comparative predictions. Those comparative predictions can be used to evaluate the differences in incidence curves between genotypes or between alternative carcinogenic environments.

Chapter 8 uses the theory to evaluate shifts in incidence curves between individuals who inherit distinct predisposing genotypes. I begin by placing two classical comparisons between inherited and noninherited cancer within my quantitative framework. The studies of Ashley (1969a) on colon cancer and Knudson (1971) on retinoblastoma made the appropriate comparison within the multistage framework, demonstrating that the inherited cases were born one stage advanced relative to the noninherited cases. I show how to make such quantitative comparisons more simply and to evaluate such comparisons more rigorously, easing the way for more such quantitative comparisons in the evaluation of cancer genetics. Currently, most research compares genotypes only in a qualitative way, ignoring the essential information about rates of progression.

I continue Chapter 8 by applying my framework for comparisons between genotypes to data on incidence in laboratory populations of mice. In one particular study, the mice had different genotypes for mismatch repair of DNA lesions. I show how to set up and test a simple comparative hypothesis about the relative incidence rates of various genotypes in relation to predictions about how aberrant DNA repair affects progression. This analysis provides a guide for the quantitative study of rates of progression in laboratory experiments. I finish this chapter with a comparison of breast cancer incidence between groups that may differ in many predisposing genes, each of small effect. Such polygenic inheritance may explain much of the variation in cancer predisposition. I develop the quantitative predictions of incidence that follow from the theory, and show how to make appropriate comparative tests between groups that may have relatively high or low polygenic predisposition. The existing genetic data remain crude at present. But new genomic technologies will provide rapid increases in information about predisposing genetics. My quantitative approach sets the framework within which one can evaluate the data that will soon arrive.

Chapter 9 compares incidence between different levels of chemical carcinogen exposure. Chemical carcinogens add to genetics a second major way in which to test comparative predictions about incidence in response to perturbations in the underlying mechanisms of progression. I first discuss the observation that incidence rises more rapidly with duration of exposure to a carcinogen than with dosage. I focus on the example of smoking, in which incidence rises with about the fifth power of the number of years of smoking and about the second power of the number of cigarettes smoked. This distinction between duration and dosage, which arises in studies of other carcinogens, sets a classic puzzle in cancer research. I provide a detailed evaluation of several alternative hypotheses. Along the way, I develop new quantitative analyses to evaluate the alternatives and facilitate future tests.

The next part of Chapter 9 develops the second classic problem in chemical carcinogenesis, the pattern of incidence after the cessation of carcinogen exposure. In particular, lung cancer incidence of continuing cigarette smokers increases with approximately the fifth power of the duration of smoking, whereas incidence among those who quit remains relatively flat after the age of cessation. I provide a quantitative analysis of alternative explanations. Finally, I argue that laboratory studies can be particularly useful in the analysis of mechanisms and rates of progression if they combine alternative genotypes with varying exposure to chemical carcinogens. Genetics and carcinogens provide different ways of uncovering failure and therefore different ways of revealing mechanism. I describe a series of hypotheses and potential tests that combine genetics and carcinogens.

Chapter 10 analyzes age-specific incidence for the leading causes of death. I evaluate the incidence curves for mortality in light of the multistage theories for cancer progression. This broad context leads to a general multicomponent reliability model of age-specific disease. I propose two quantitative hypotheses from multistage theory to explain the mortality patterns. I conclude that multistage reliability models will develop into a useful tool for studies of mortality and aging.

Part III discusses evolutionary problems. Cancer progresses by the accumulation of heritable change in cell lineages: the accumulation of heritable change in lineages is evolutionary change.

Heritable variants trace their origin back to an ancestral cell. If the ancestral cell of a variant came before the most recent zygote, then the individual inherited that variant through the parental germline. The frequency of inherited variants depends on mutation, selection, and the other processes of population genetics. If the ancestral cell of a variant came within the same individual, after the zygote, then the mutation arose somatically. Somatic variants drive progression within an individual.

Chapter 11 focuses on germline variants that determine the inherited predisposition to cancer. I first review the many different kinds of inherited variation, and how each kind of variation affects incidence. Variation may, for example, be classified by its effect on a single locus, grouping together all variants that cause loss of function into a single class. Or variation may be measured at particular sites in the DNA sequence, allowing greater resolution with regard to the origin of variants, their effects, and their fluctuations in frequency. With resolution per site, one can also evaluate the interaction between variants at different sites. I then turn around the causal pathway: the phenotype of a variant—progression and incidence—influences the rate at which that variant increases or decreases within the population. The limited data appear to match expectations: variants that cause a strong shift of incidence to earlier ages occur at low frequency; variants that only sometimes lead to disease occur more frequently.

I finish Chapter 11 by addressing a central question of biomedical genetics: Does inherited disease arise mostly from few variants that occur at relatively high frequency in populations or from many variants that each occur at relatively low frequency? Inheritance of cancer provides the best opportunity for progress on this key question.

Chapter 12 focuses on somatic variants. Mitotic rate drives the origin of new variants and the relative risk of cancer in different tissues. For example, epithelial tissues often renew throughout life; about 80–90% of human cancers arise in epithelia. The shape of somatic cell lineages in renewing tissues affects how variants accumulate over time. Rare stem cells divide occasionally, each division giving rise on average to one replacement stem cell for future renewal and to one transit cell. The transit cell undergoes multiple rounds of division to produce the various short-lived, differentiated cells. Each transit lineage soon dies out; only the stem lineage remains over time to accumulate heritable variants. I review the stem-transit architecture of cell lineages in blood formation (hematopoiesis), gastrointestinal and epidermal renewal, and in sex-specific tissues such as the sperm, breast, and prostate.

I finish Chapter 12 by analyzing stem cells divisions and the origin of heritable variations. In some cases, stem cells divide asymmetrically, one daughter determined to be the replacement stem cell, and the other determined to be the progenitor of the short-lived transit lineage. New heritable variants survive only if they segregate to the daughter stem cell. Recent studies show that some stem cells segregate old DNA template strands to the daughter stem cells and newly made DNA copies to the transit lineage. Most replication errors probably arise on the new copies, so asymmetric division may segregate new mutations to the short-lived transit lineage. This strategy reduces the mutation rate in the long-lived stem lineage, a mechanism to protect against increased disease with age.

Chapter 13 analyzes different shapes of cell lineages with regard to the accumulation of heritable change and progression to cancer. In development, cell lineages expand exponentially to produce the cells that initially seed a tissue. By contrast, once the tissue has developed, each new mutation usually remains confined to the localized area of the tissue that descends directly from the mutated cell. Because mutations during development carry forward to many more cells than mutations during renewal, a significant fraction of cancer risk may be determined in the short period of development early in life. Once the tissue forms and tissue renewal begins, the particular architecture of the stem-transit lineages affects the accumulation of heritable variants. I analyze various stem-transit architectures and their consequences. Finally, I discuss how multiple stem cells sometimes coexist in a local pool to renew the local patch of tissue. The long-term competition and survival of stem cells in a local pool determine the lineal descent and survival of heritable variants.

Chapter 14 describes empirical methods to study cell lineages and the accumulation of heritable change. Ideally, one would measure heritable diversity among a population of cells and reconstruct the cell lineage (phylogenetic) history. Historical reconstruction estimates, for each variant shared by two cells, the number of cell divisions back to the common ancestral cell in which the variant originated. Current studies do not achieve such resolution, but do hint at what will soon come with advancing genomic technology. The current studies typically measure variation in a relatively rapidly changing aspect of the genome, such as DNA methylation or length changes in highly repeated DNA regions. Such studies of variation have provided insight into the lineage history of clonal succession in colorectal stem cell pools and the hierarchy of tissue renewal in hair follicles. Another study has indicated that greater diversity among lineages within a precancerous lesion correlate with a higher probability of subsequent progression to malignancy.

I finish Chapter 14 with a discussion of somatic mosaicism, in which distinct populations of cells carry different heritable variants. Mosaic patches may arise by a mutation during development or by a mutation in the adult that spreads by clonal expansion. Mosaic patches sometimes form a field with an increased risk of cancer progression, in which multiple independent tumors may develop. Advancing genomic technology will soon allow much more refined measures of genetic and epigenetic mosaicism. Those measures will provide a window onto cell lineage history with regard to the accumulation of heritable change—the ultimate explanation of somatic evolution and progression to disease.

Chapter 15 summarizes and draws conclusions.