SLAS Discovery 29 (2024) 100194

Contents lists available at ScienceDirect

SLAS Discovery

journal homepage: www.elsevier.com/locate/slasd

The history, landscape, and outlook of human cell line authentication and security

Elijah Harbut [a,b,c], Yiorgos Makris [d], Alexander Pertsemlidis [e,f,g], Leonidas Bleris [a,b,c,*]

[a] Department of Bioengineering, The University of Texas at Dallas, Richardson, TX, USA
[b] Center for Systems Biology, The University of Texas at Dallas, Richardson, TX, USA
[c] Department of Biological Sciences, The University of Texas at Dallas, Richardson, TX, USA
[d] Department of Electrical and Computer Engineering, The University of Texas at Dallas, Richardson, TX, USA
[e] Department of Pediatrics, The University of Texas Health Science Center at San Antonio, San Antonio, TX, USA
[f] Department of Cell Systems & Anatomy, The University of Texas Health Science Center at San Antonio, San Antonio, TX, USA
[g] Greehey Children's Cancer Research Institute, The University of Texas Health Science Center at San Antonio, San Antonio, TX, USA

1. Introduction

The phenomenon of cellular cross-contamination and misidentification ("CCCM") was identified more than six decades ago and continues to pose significant challenges for laboratories globally. The prevalence of CCCM is alarmingly high, with estimates suggesting that up to 30% of cell lines may be misidentified, leading to substantial financial losses [2–4]. Accelerating progress in biotechnology has intensified the complexity of ensuring experimental reproducibility and data validity and compounds the need to protect the intellectual property that arises from cell line engineering. This calls for a unified approach from the scientific community—including researchers, journal editors, and funding bodies—to adopt new technologies and establish robust cellular authentication protocols as standard practice in life science research.

Despite fifty-five years of attention through reviews, letters, calls to action, surveys, and technical innovations [5–8], cellular cross-contamination and misidentification persist, with global rates estimated to be between 14% and 46%. The International Journal of Cancer ("IJC"), an early proponent of mandated authentication for publication, observed cell line problems in approximately a fifth of submitted manuscripts over the three-year period from 2018 to 2021 [4, 9–11]. The propagation of genetically engineered sublineages exacerbates this problem, highlighting the urgent need for widespread adoption of authentication protocols.

Like CCCM, mycoplasma contamination dates to the dawn of cell culture and undermines the validity and reproducibility of experimental results. First reported in 1956, many cell cultures and established cell lines were found to be contaminated with mycoplasma. The small size and atypical antibiotic resistance of mycoplasma enable it to pass through filters and multiply in media prodigiously [12]. Mycoplasma contamination rarely induces morphological changes, instead altering biosynthetic pathways, genetics, and general cell function, and therefore must be assayed chemically or biologically [13–15]. In contrast to its response to CCCM, the scientific community took swift and decisive action to minimize the potential damage posed by mycoplasma contamination. Only a few years after the first reported identification of contamination, the United States Public Health Service ("USPHS") required a test for cell cultures used to produce viral vaccines. A decade later, the Food and Drug Administration ("FDA") codified guidance for mycoplasma contamination testing in 21 CFR 610.30. These swift actions led to a significant reduction in primary cell line contamination and secondary contamination [9,14].

Previous reviews have highlighted authentication methods and databases of cell fingerprints as potential mechanisms by which the threat of CCCM could be mitigated [10,16–18]. Short tandem repeats ("STRs") are microsatellites with repeat units 2 to 7 base pairs in length, with the number of repeats varying among cell lines, making them effective for cell line authentication. Adopted by pharmaceutical companies and academia for the characterization of primary and immortalized cells [11,19], STR profiling is well established with standard protocols published by the American National Standards Institute ("ANSI") [20–22], relatively inexpensive, and flexible – its resolution can be improved by increasing the number of profiled loci [23,24].

Immortalized cell lines have revolutionized molecular biology by simplifying research processes and reducing costs, allowing cultures to be duplicated, cryogenically preserved for long periods, revived, and

"If we can't tell one cell from another, we have grave problems." – Robert H. Bassin, PhD [1]

* Corresponding author at: Department of Bioengineering, The University of Texas at Dallas, Richardson, TX, USA.
E-mail address: bleris@utdallas.edu (L. Bleris).

https://doi.org/10.1016/j.slasd.2024.100194
Received 7 June 2024; Received in revised form 30 September 2024; Accepted 4 November 2024
Available online 8 November 2024

immediately employed in experiments. Immortalized cell lines typically retain similarity to their progenitors in culture, including some level of clonal heterogeneity, but they develop genetic instability and frequent aneuploidy [25–27]. Genetic drift can be significant enough over extended passaging to render a line unrepresentative of its primary source. A further drawback of this phenomenon is that CCCM or other issues with a line may be attributed to qualities acquired through genetic de-repression. For instance, irregularities in Chang liver cells were initially ascribed to this, but the cells were eventually confirmed to be HeLa-derived [2,22,28–30]. Unlike immortalized cells, primary cells are only able to divide a fixed number of times. They remain the gold standard for therapeutic tests because of genotypic and phenotypic fidelity to their source tissues but are still susceptible to CCCM [27].

The mechanisms for CCCM are diverse, ranging from use of the same media and reagents to culture different cell lines to mislabeling of cultures and laboratory disorganization [27,31]. Historically, intraspecies contamination is more common since interspecies contamination is more easily detected [27]. A striking example is the STR-based detection of HeLa contamination in the human endometrial cell line HES, highlighting the risk of secondary contamination [32,33]. This phenomenon raises questions about the existence of authentic stock for any given cell line and emphasizes the importance of transparency and self-reporting within the scientific community.

Despite the availability of tools to prevent CCCM, a comprehensive and mandated approach akin to the one taken against mycoplasma contamination is lacking. The reluctance to address CCCM is often due to a misconception among principal investigators who do not perceive their labs as susceptible. However, contamination rates and the frequency of high-profile cases demonstrate the universal vulnerability to CCCM [34–37]. Establishing robust protocols and regular authentication practices can substantially reduce the incidence of CCCM. The cost of these preventive measures is minimal compared to the potential financial and reputational damage of compromised research.

The consequences of CCCM are dire, leading to invalid data, irreproducible experiments, and the wasteful expenditure of valuable resources. Scientific reproducibility is intertwined with data integrity, a fact highlighted by multiple assessments of influential research papers, which revealed a crisis of reproducibility. Amgen attempted to confirm 53 landmark oncology papers in the early 2000s and successfully validated 6 (11%). Bayer Healthcare released a similar report disclosing that they were only able to verify 25% of preclinical research [38,39]. The inability to replicate findings from a significant percentage of studies indicates a systemic issue that extends beyond individual lapses in methodology. Surveys have indicated that 70% of scientists have tried and failed to reproduce manuscript findings [40–42]. A 2014 National Academy of Sciences committee planning to emphasize the necessity of transparency and reproducibility selected 193 experiments from 53 important cancer biology papers published from 2010–2012, corresponded with the authors, and undertook the replication of the findings [43]. The "Reproducibility Project: Cancer Biology" was unable to find sufficient protocol information for reproducing the experiments in any of the original manuscripts, and even when they were able to rerun experiments, the reconstructed findings were often of greater variability or lower efficacy than the original ones [43–46]. These conclusions do not negate the entire profiled body of work, but they do underscore the need for improved communication and transparency in experimental reporting.

2. Historical context

The history of cellular cross-contamination and misidentification is punctuated by significant events that have shaped the current understanding and approach to this persistent issue (Figs. 1,2). The harvest and culture of Henriette Lacks' cervical adenocarcinoma cells by Dr. George Gey in 1951 [3,47,48] represented the first "cancer in a test tube" and catalyzed an era of research based on immortalized human cells. Since then, immortalized cells have enabled a rapid expansion of the applications and interrogative abilities of in vitro science. For

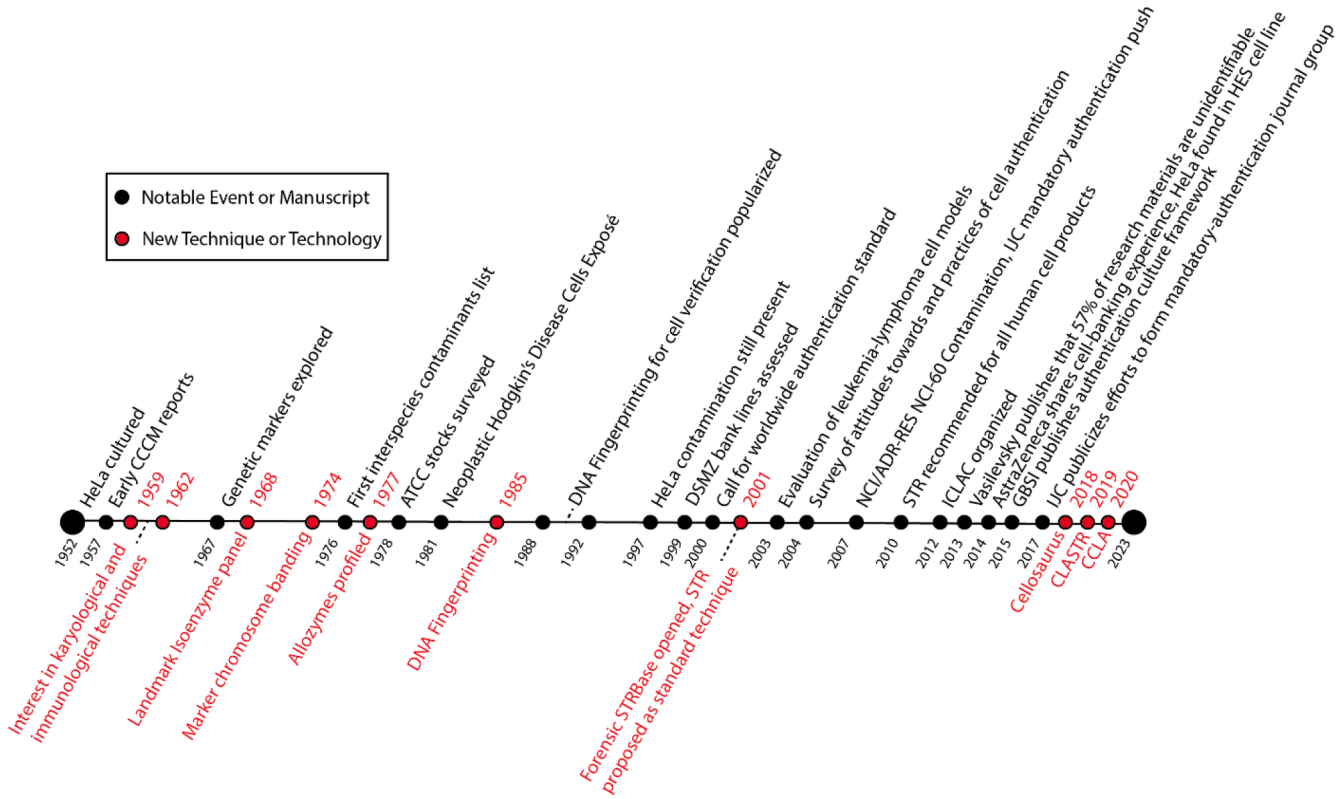

Fig. 1. Historical Timeline. Significant events in cellular cross-contamination and misidentification.

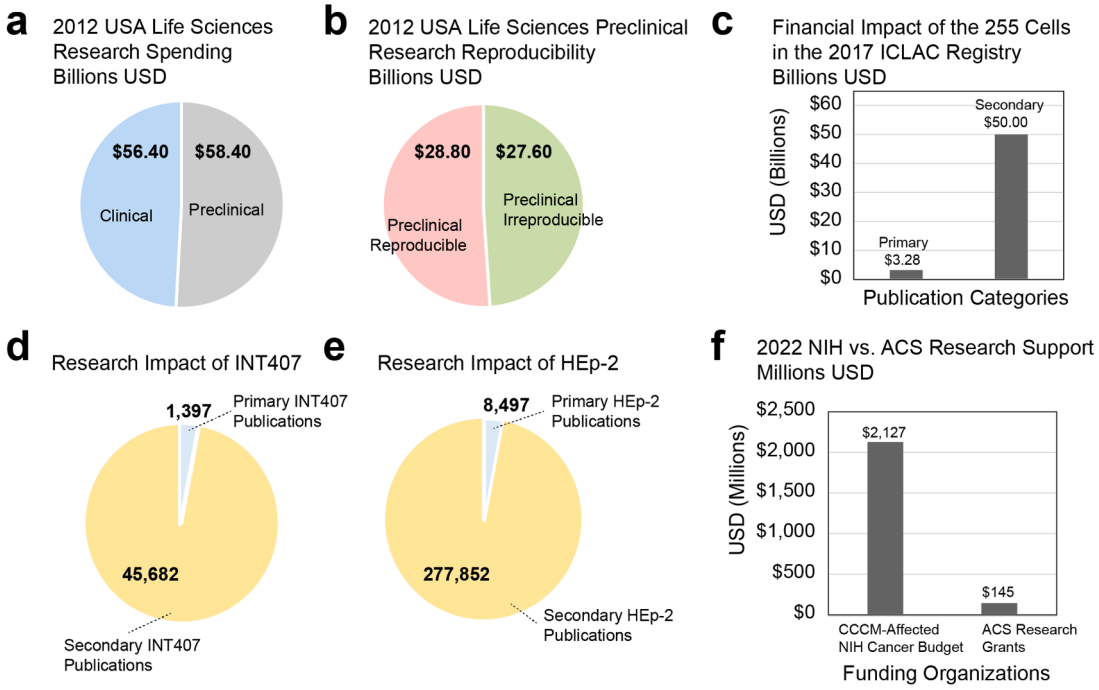

Fig. 2. Life Sciences Spending, Reproducibility, and CCCM. (**A**) In 2012, life sciences spending in the USA was split between preclinical and clinical research [99]. (**B**) More than half of the research produced from 2012 American life sciences spending was irreproducible [99]. (**C**) Work based on biology research tainted by CCCM demonstrates the impact of the issue [3,4,100]. (**D**) INT407, like other misidentified lines, has a major impact on secondary publications [3]. (**E**) HEp-2 [HeLa] research is still frequently cited [3]. (**F**) The annual financial effect of CCCM is more than 15 times the overall funding of research by the American Cancer Society [87,102].

example, human cells are ubiquitously used in the production of vaccines, including those for COVID-19 [49].

A notable early example is the work of Jonas Salk in the 1950s. While developing his polio vaccine, Salk utilized various cell types, including HeLa cells [50]. It was during a series of clinical trials that Salk, aiming to induce immune responses in cancer patients, inadvertently injected them with HeLa cells instead of the intended cynomolgus monkey heart cells. This mishap, later confirmed by Dr. Nelson-Rees, a cell culture expert and early proponent of cell line authentication [51], was significant enough that, had the same oversight occurred today, the polio vaccine would not meet FDA approval standards [14,52]. Dr. Salk's experience was far from unique. Dubbed by some "the ghosts of HeLa", these cells have been a notorious contaminant of cell culture work for decades [3,47,53,54].

Early cases of suspected CCCM were communicated in 1957 by McCulloch and Westwood, who studied the MCN and ERK-1 cell lines, respectively [31]. These reports motivated interest in karyological and immunological characterization techniques by Rothfels (1959), Defendi (1960), and Brand and Syverton (1962), in addition to the formation of the Cell Culture Collection Committee in 1960 [55–58]. Walter Nelson-Rees' "Cell Culture Lab" was created and funded by the National Cancer Institute ("NCI") in 1961 [59]. Stanley Gartler explored the potential of genetic markers for CCCM detection [60] and examined HeLa contamination in a landmark isoenzyme panel of popular cell lines [54].

Through the 1970s, the field witnessed efforts to combat CCCM, such as the Animal Cell Culture Collection [55], but the response from the scientific community was tepid. Between 1970 and 1974, Walter Nelson-Rees wrote 25 papers on CCCM, yet there was little reaction from the field [61]. As CCCM continued to spread, Nelson-Rees developed a technique for marker chromosome banding [62]. He published the first list of interspecies contaminants in 1976 and analyzed recently established lines [59,63]. The range of possible CCCM detection methods was expanded by O'Brien, et al. with the introduction of enzymatic polymorphisms (allozymes) in 1977 [30]. Following this, Jonas Salk spoke at the 1978 Lake Placid Conference on the Cellular Production of Vaccines and mentioned his own experiences with CCCM; these were omitted from the official record by the editors of the conference's proceedings [59,64]. In 1978, Lavappa produced a survey of American Type Culture Collection ("ATCC") stocks for HeLa contamination [65]. A year later, in another prominent case of CCCM, an anonymous reviewer requested that Todd and Furcinitti remove details surrounding suspected cellular contamination from a manuscript [59,66,67]. This situation was brought to light after Nelson-Rees identified the cells used by the duo as HeLa cells [67]. They commented that "not all attempts by scientists to be honest and thorough are accommodated by journal editorial policies" [59].

High-profile exposés continued to appear, driven largely by Nelson-Rees. Three labs independently evaluated "neoplastic Hodgkin's disease cells" produced by Harvard pathologist John C. Long and revealed them to be non-Hodgkin's cells of human and Colombian brown-footed owl monkey origins [68]. This example of misconduct marked the first time the Proceedings of the National Academy of Sciences ("PNAS") had to "disown a published paper" and earned the ire of reporter David Dickson, who called it a "corruption of … the literature" [69]. John Maddox, then-editor of *Nature*, accused Nelson-Rees of vigilantism, saying there was "no reason to suppose the few cases of dishonesty that have come to light are in any sense the tip of the iceberg" [59]. In contrast, former Stanford president Donald Kennedy asserted that "whatever one sees is probably only a small fraction of what actually goes on" [59]. Nelson-Rees revised and expanded his list of contaminated lines in the same year [7]. Responses to his work ranged from gentle retractions of cell lines (as in the cases of Bassin, Essex, and Kanki) to "unsigned telegrams containing one-way tickets to and promises of positions in Uganda" [59].

The invention of deoxyribonucleic acid ("DNA") fingerprinting by Jeffreys, et al. in 1985 was quickly applied to cell line recognition [61, 70–73], with subsequent discoveries of HeLa contamination in various lines underscoring the extent of the CCCM problem. A few years later,

Ogura, et al. uncovered HeLa contamination in the lines JTC-3, OG, and OE [74]. In 1999, scientists at the Deutsche Sammlung von Mikroorganismen und Zellkulturen ("DSMZ") profiled 252 human cell lines submitted by cell repositories and originating laboratories using mini and microsatellite techniques, calculating that 18% of the donated lines and 29% of the 93 sources were affected by CCCM [6]. The authors discussed CCCM in relation to hematopoietic lines, and later commentaries on the work suspected higher "true percentages" of CCCM because secondarily acquired lines have generally exhibited greater CCCM rates [24,75]. The 1990s and 2000s saw many other instances of CCCM and the proliferation of DNA fingerprinting methods, such as amplified fragment length polymorphism ("AmpFLP"), restriction fragment length polymorphism ("RFLP"), and STR profiling [31,76–79].

By the end of the 20th century, with the establishment of cell line repositories, the scope of CCCM was becoming clearer. The new millennium brought with it studies showing that a significant percentage of cell lines were affected, hinting at an even more widespread issue than previously thought and generating a collective demand for mandatory authentication requirements for publication. In 2000, a team of American, British, and German researchers signed a resolution demanding a worldwide mandatory authentication requirement for publishing after the DSMZ and ATCC reported on the extent of CCCM in cell line repositories [80]. In 2001, the first online database of STR markers, STRBase, was released to the forensics community. STR fingerprinting was proposed as an "international reference standard" that same year [22,81]. These were milestones in establishing a systematic approach to cell line authentication, but despite these advancements, surveys revealed a lackadaisical attitude towards CCCM, with many researchers failing to verify their cell lines regularly. In 2003, Drexler, et al. fingerprinted 550 leukemia-lymphoma cells and calculated a 29.7% rate of CCCM [82]. In 2004, UC Berkeley's Gertrude Buehring presented a survey of 483 researchers meant to probe attitudes towards CCCM and mitigation practices. Fewer than 50% of respondents regularly verified lines using standard techniques and previously uncovered HeLa contaminants were being used by 9% of scientists [83]. 2005 saw the publicizing of contamination in the putative normal colon epithelial cell line and an article by Masters quantifying the continued use of the recognized contaminated cell lines INT407, WISH, Chang Liver, HEp-2, and KB [37,61]. The issue of CCCM was reviewed two years later, revealing an 18–36% CCCM rate. The NCI disclosed that a contaminated cell line (then called MF-7/ADR and now called NCI/ADR-RES) had infiltrated its NCI-60 Human Tumor Cell Lines Screen [75,84].

In response to the mounting evidence and continued spread of CCCM, the International Journal of Cancer proposed that a consortium of journals adopt mandatory authentication policies [4]. Calls for a standardized method of cellular authentication mounted once more in 2009 [16,47,85]. Notably, one group of authors emphasized the lack of action against CCCM by reiterating that "KB is not an oral squamous carcinoma cell line", despite this having been discovered in 1978 [65,86]. 2010 saw Capes-Davis, Dirks, and others recommending STR authentication for "human cell lines, stem cells, and tissues" [52].

The following years saw the organization of the International Cell Line Authentication Committee ("ICLAC") and the release of databases and tools such as Cellosaurus and CLASTR, which aimed to provide resources for accurate cell line identification. ICLAC was organized in 2012 after a collection of scientists formulated a standard using STR profiling (ASN-0002) [10,87]. In 2013, Vasilevsky, et al. evaluated 200 papers and reported that 57% of cell lines were not uniquely identifiable; the true incidence was probably much higher because the survey did not include "cells displaced by unknown cell types, incorrect first culture identifications, and individual lab data" [88]. In 2014, AstraZeneca shared the experiences of its scientists in trying to establish a centralized corporate cell bank, outlining guidelines for "cells used in biomedical research" [19,76]. CCCM in the HES cell line was detected; public announcements and interlaboratory testing followed [32]. In 2015, the Cancer Cell Authentication and Standards Task Force of the Global Biological Standards Institute published a framework for changing the culture surrounding cellular authentication [10,89]. In 2017, the IJC publicized a case study describing their experience mandating cell line verification for published work [4]. The Cellosaurus database went online in 2018 and its search tool, CLASTR, was released in 2019 [90,91]. The RNA-Seq cell line authentication server CCLA was released in 2020 [92].

This historical journey, from the early days of HeLa contamination to the establishment of comprehensive databases and standards, underscores the evolving battle against CCCM. It is in some ways a testament of the failure of the scientific community to respond comprehensively to persistent challenges to research integrity.

3. Current challenges and issues

The persistent threat of cellular cross-contamination and misidentification continues to cast a long shadow over the landscape of contemporary scientific research. Despite over six decades of awareness and significant technological advances, the prevalence of CCCM remains disturbingly high, with estimates suggesting that 14% to 46% of cell lines are misidentified. This widespread issue not only undermines the validity of research findings but also leads to substantial economic losses, with several billion dollars estimated to be wasted annually due to compromised studies.

The advent of biotechnology and the acceleration of cellular engineering have further complicated the CCCM challenge. As researchers push the boundaries of science, particularly in the fields of genetics and cellular biology, the need for accurate cell line authentication becomes increasingly critical. The safeguarding of intellectual property, a byproduct of cell line engineering, adds another layer of complexity to the issue, making it imperative for researchers to verify the lineage and purity of their cell lines to protect their work and contributions to the field.

One of the historical parallels to the CCCM issue is mycoplasma contamination, which has plagued cell cultures since the early days of cell culture techniques. The insidious nature of mycoplasma, with its small size and resistance to common antibiotics, has necessitated the development of specialized assays for detection. In stark contrast to the slow response to CCCM, the scientific community responded decisively to the mycoplasma threat, leading to stringent testing requirements and a significant reduction in contamination rates. This swift action against mycoplasma serves as a benchmark for the level of response required to mitigate CCCM.

Moreover, the propagation of genetically engineered cell lines has introduced a new dimension to the CCCM issue. With these advanced cell lines comes the risk of further complicating the already challenging task of maintaining pure cultures. Instances of secondary contamination, such as the notable case of HeLa contamination in the human endometrial cell line HES, exemplify the intricate web of contamination that researchers must navigate. The problem is compounded by the fact that contamination can occur at any stage, from cell line derivation to laboratory handling, making the tracing and verification of cell line authenticity a daunting task.

Despite the availability of sophisticated authentication methods like Short Tandem Repeat profiling, there remains a gap in universal adoption and implementation of such protocols. The resistance to addressing CCCM proactively can be attributed to various factors, including a lack of awareness among principal investigators about the potential risks to their laboratories and the misconception that non-tissue-specific experiments are exempt from the threat of contamination. This complacency is at odds with the high-profile cases and contamination rates, which unequivocally demonstrate that no laboratory is immune.

To combat these challenges, it is essential to establish protocols that are as fundamental to scientific research as maintaining a laboratory notebook. If research teams were to adopt cell line authentication at the onset and conclusion of each experimental series, and periodically

thereafter, the incidence of CCCM could be significantly reduced. The cost of authentication, often cited as a barrier, pales in comparison to the economic and reputational damage that can result from compromised research.

The path to addressing current challenges and issues in CCCM lies in a concerted effort from all stakeholders in the scientific community. It involves not only the adoption of established authentication protocols but also a cultural shift towards transparency and accountability in research practices. The future of scientific integrity and reproducibility hinges on our ability to confront and overcome the challenges posed by CCCM.

4. Technological and methodological advances

Academic and commercial entities have developed a diverse suite of technologies and services in the fight against cellular cross-contamination and misidentification, but challenges remain, particularly with intraspecies contamination [11,93–96]; Tables 1,2.

4.1. Conventional karyotyping

Conventional karyotyping has been reported to detect contamination when present in as little as 1% of a culture [31,97]. It requires substantial time and expertise, making newer approaches more attractive [13,24].

Table 1
Historic, popular, and new techniques for CCCM detection and cellular authentication.

Techniques for CCCM Detection and Cellular Authentication		
Name	General Workflow	Refs.
Karyotyping		
Classical Karyotyping (CK)	Fix cells in metaphase, arrange, analyze	[11,28,37,85,116,126]
G-Banding (Giemsa Staining)	CK with staining	[11,44,49]
Spectral Karyotyping (SKY, FISH)	CK with fluorescent probe hybridization	[11,49,85]
Antigen-Antibody Reactions		
Fluorescent Antibody Staining	Aliquot cells, treat, analyze	[11,63,117,118]
HL-A Typing	Aliquot cells, treat, analyze	[11]
Enzyme Analysis		
Isoenzyme Analysis	Extract enzymes, gel electrophoresis, analyze	[11,19,37,67,85]
Allozyme Analysis	Extract enzymes, gel electrophoresis, analyze	[32,44,75]
DNA Fingerprinting		
*All methods involve gDNA extraction and PCR amplification		
Vanilla Fingerprinting	Southern blot, analyze	[50,52,78,85]
RFLP	Digestion, Southern blot, analyze	[50,85,127]
AmpFLP	Digestion, adapter ligation, 2*PCR, gel electrophoresis, analyze	[50,51,85,127]
STR	Capillary electrophoresis, analyze	[25,34,105,120,126,128]
SNP	Capillary electrophoresis, analyze	[25,105,120,123,125]
mtDNA	Sequence, analyze	[121,122,126]
Cell Morphology		
Tracking Traits	Investigation-dependent	[112,119,126]
New Techniques		
Aggregation-Induced Emission	Treat cells with nanomaterials, analyze	[130]
Cytotoxicity	Treat, analyze	[129]
RNA-Seq	Extract RNA, prepare library, sequence, analyze	[86,131]
NGS	Extract DNA, prepare library, sequence, analyze	[138]
PUF	Sequence, generate PUF matrix	[139]
MALDI-ToF Mass Spectrometry	Prepare cells, mass spectroscopy, analyze	[23]
Neural Network Image Analysis	Collect images, analyze with model	[140,141]

4.2. G-banding (Giemsa Staining)

Digesting cellular DNA with trypsin and applying Giemsa stain results in unique monochromatic patterns that vary in band density, value, and location [30,35]. It is one of the most popular staining techniques. Highly similar banding patterns may indicate a shared origin but different lineages [31].

4.3. Spectral karyotyping ("SKY")

Gene-specific fluorescent "SKY" probes are bound to chromosomes, creating a polychromatic karyotype. This simplifies visual analyses [37]. SKY is an advanced form of fluorescent in situ hybridization ("FISH") and can be used for determination of ploidy, detection of difficult marker chromosomes, and identification of genetic changes [31,35,78].

4.4. Antibody-antigen reactions and fluorescent antibody staining

Antigen-antibody reactions can only positively detect cells, but they are sensitive to contamination $\geq$ 0.1% of a culture [31,78]. Antibodies targeting species-specific antigens are attached to probes for fluorescent staining [55,98,99]. HL-A typing is also popular. It has been historically recommended these methods be used with other technologies for interspecies tests because they may yield false negatives [31].

4.5. Isoenzyme and allozyme analysis

The electrophoretic mobilities of chosen enzymes from a given cell are used to generate a unique gel banding pattern that can generally detect contamination affecting $\geq$ 5–10% of the culture population [24,31,78,97]. Enzymes can be further described through colorimetric staining, pH preferences, substrate specificities, inhibitor responses, and thermostabilities [31,78]. Allozyme analysis focuses on the alleles of a single enzymatic gene, whereas isoenzyme analysis targets different enzymatic genes that perform the same task [18,30,68].

4.6. DNA fingerprinting (RFLP, AmpFLP, SNP, barcodes, STR)

DNA hybridization techniques have been applied in the forensic and biological sciences since 1985 [71,100]. Early fingerprinting targeted 15 base pair polymorphic regions called minisatellites or variable number of tandem repeats ("VNTRs"), while current techniques target STRs or microsatellites where the repeat unit is typically 5 or fewer base pairs [78,100]. DNA Fingerprinting may have difficulty discriminating between closely-related cells, but it can detect contamination between human cell lines when it composes $\geq$ 5% of a culture [31,97]. Alternatives surveying single nucleotide polymorphisms ("SNPs") (detects contamination $\geq$ 3%), the cytochrome *c* oxidase I gene (reveals contamination $\geq$ 1%), and other selected genes have been explored, but they have not gained traction because of inherent sampling volume requirements and intraspecies limitations [97,101–108]. RFLP, AmpFLP, and STR profiling have all been popular, but high-loci STR assays have prevailed as the current institutional technique [21,36,78,79,109,110]. The American National Standards Institute based its current authentication protocol, ANSI-0002, on STR because of the low cost, high discrimination power, and customizability. Many of the traditional limits of STR-based methods are solved by increasing the number of sampled loci [10,111].

4.7. Machine learning

Machine learning approaches to cell characterization may rely on data generated from microfluidics systems, labeling technologies (fluorescent probes and chemical stains), or a variety of microscopy

Table 2
Examples of purchasable STR Solutions. STR is the authentication standard; purchasable STR solutions are listed to ease adoption.

Purchasable STR Solutions				
Technique	Organization	External Academic Cost (Per Sample)	Academic, For-Profit, Nonprofit	Category (In-House Kit, Mail Service, Machine)
STR (18 Loci)	ATCC	$203.00	Nonprofit	Mail Service
STR (Setting Up In-Lab)	Promega	$66,000-$72,000*	For-Profit	Capillary Electrophoresis Machine (Spectrum Compact CE)*
STR (Setting Up In-Lab)	Promega	$3.50	For-Profit	Capillary Electrophoresis Machine Consumables
CLA GlobalFiler™ PCR Amplification Kit	ThermoFisher	$6,270.00	For-Profit	Kit
Cell Line Authentication Barcodes	Eurofins	$108.36	For-Profit	Mail Service
STR (24 Loci)	The University of Utah	$91.50	Academic	Mail Service
STR (16 Loci)	ECACC	$159.81	Nonprofit	Mail Service
STR (9 Loci)	The University of Illinois at Urbana-Champaign	$65.00	Academic	Mail Service
STR (18 Loci)	Johns Hopkins	$140.00	Academic	Mail Service
STR (16 Loci)	Northwestern University	$106.25	Academic	Mail Service
STR (16 Loci)	The University of Arizona	$66.00	Academic	Mail Service

* Cost for the Capillary Electrophoresis Machine, rather than per sample

techniques [112–115]. Algorithms are tasked with cell "segmentation, tracking, and classification" [116,117]. The computational identification and sorting of cells has been used to distinguish between cancerous and noncancerous cells, stem cells, types of red and white blood cells, and cells of different states [118–123]. Tracking cell doubling times or growth rates, passage numbers, phenotypes, and other properties can give insight into whether a culture is impacted by CCCM [101,124]. As evidenced by these use-cases, machine learning methods may be potentially helpful for CCCM-related applications.

4.8. Other techniques

Chemical approaches to cell authentication involving aggregation-induced emission and the comparison of IC_{50} cytotoxicity values have been recently explored [125,126]. Expression profiling by RNA-seq has been able to correctly identify cell lines and estimate cross-contamination [127,128]. Research Resource Identifiers ("RRIDs") have been shown to decrease CCCM [88,129,130]. Combined with recently developed intra-lab cell versioning software and accepted verification technologies (i.e., STR profiling), RRIDs could potentially improve tracking of a given line's usage [131–133]. Intact cell MALDI-ToF mass spectrometry—normally used for the clinical recognition of bacteria—has been successfully harnessed to distinguish drug-adapted sublineages from their parental lines [8]. Authentication experiments with Next-Generation Sequencing ("NGS") have shown extremely high accuracy, even in lines derived from the same parent [134].

4.9. Genome editing

Genome editing innovations primarily driven by CRISPR technologies [135–140] are driving a paradigm shift in molecular bioscience. Engineered cell lines are now easily produced from a single parent clone, increasing the need for methods for provenance attestation and the protection of intellectual property rights [141–143]. Such cell lines are becoming a significant portion of the bioeconomy, with the global economic impact of bio-based products, services, and processes predicted to reach up to $4 trillion annually [144].

This rapid development is similar to how the advent of the integrated circuit ("IC") drove an explosion of applications for computing. There, counterfeit versions created a need for source attestation [145], which was addressed through the discovery of physical unclonable functions ("PUFs"). Silicon PUFs exploit the inherent stochasticity in semiconductor manufacturing [146]. Similarly, biological PUFs exploit genetic differences in the process of engineering cell lines to provide robust, unique, and irreproducible identifying signatures of cells [147]. PUFs can decrease the accidental use of misidentified or counterfeit cell lines and make the intentional production of counterfeit lines more difficult, reducing CCCM, protecting consumers by ensuring cell line quality, and securing commercial cell products against illegal duplication and sale. Technologies such as STR profiling, karyotyping, and barcoding do not meet the needs of fledgling procurement networks because they do not exhibit the three requisite attributes for source attestation: robustness, uniqueness, and unclonability [134].

5. Economic and global impact

The ramifications of CCCM extend beyond the scientific community, and include wasted resources and questionable validity of numerous studies. A stark illustration of this economic burden is the revelation that the National Institutes of Health ("NIH") may have allocated approximately $100 million to research involving the misidentified cell line NCI/ADR-RES. Given an average NIH grant of $370,000 for breast cancer research, an estimated $100 million has been spent on NCI/ADR-RES alone [10,36]. In a broader context, $114.8B was spent on life sciences research in the United States in 2012, with $56.4B going to preclinical research. Assuming a charitable 51% irreproducibility rate, Freedman, et al. estimated that at least $28B is spent annually on faulty work [148]. Even a 25% reduction of this waste could rescue billions of USD in funding [148].

The repercussions of utilizing misidentified cell lines are not trivial. In 2021, Korch and Capes-Davis evaluated the research and economic impacts of CCCM by focusing on the HEp-2 and INT407 cell lines [2]. These lines, known to be misidentified since the 1960s, were used in 8497 articles published on HEp-2 and 1397 articles on INT407. 37% of the HEp-2 articles ascribed an incorrect origin to the cell line, with the number of manuscripts growing by 250 per year. A single lab published 2.65% of this work. Although they ceased using the line after being alerted, none of their articles were retracted or labelled with alerts. The lab's publications received an average of 32.7 citations each, meaning that 45,682 papers were built upon their incorrect information [2]. The INT407 "mean citations" metric was also used to calculate that 277,852 manuscripts were impacted by HEP-2. This represents an enormous waste of resources: if it is "assumed that each article [using the two lines] was cited 15 times … the cost could be as high as $14.8 billion" [2]. Korch and Capes-Davis cited data from Horbach and Halffman's 2017 report to estimate that literature based on the 255 entries in the ICLAC Register of Misidentified Cell Lines produced a cost of $3.28

billion for primary publications and $50 billion for secondary publications [2,3,149].

The challenge of CCCM is not limited to the United States. China exemplifies the international dimension of this issue. A member of the World Federation for Culture Collections ("WFCC") since 1987 [104, 150], it has made bioeconomic development a strategic priority, promising to advance research areas that necessitate cell line usage [151,152]. With the highest annual output of scientific articles, the prevalence of CCCM in Chinese research is particularly concerning. China became the world's largest annual producer of scientific research by volume in 2016, publishing 426,000 articles that year. A 2015 STR profile of 380 samples from 113 independent labs uncovered a 25% (95/380) CCCM rate [153]. More than three quarters of the lines established in China were misidentified [153]. Two years later, a 21-STR evaluation of 278 lines originating from 28 Chinese sources found a CCCM rate of 46% (128/278); 22 of the sources had contributed contaminated lines [104]. Again, the CCCM rate among Chinese-derived cell lines was much higher: 73.2% (52/71) [104]. A 2017 review of 482 human tumor cell lines used in China showed a total 20.5% CCCM rate, the majority of which was intraspecies contamination [124]. The authors highlighted the case of the cell line ECV304, which was initially derived in 1990. ECV304 was identified as T24 human bladder carcinoma cells in 1999. Almost 70% of ECV304 publications were contributed by Chinese labs from 2004–2017 [124]. This trend is not limited to ECV304 [T24]; as Chinese-language publications using HEp-2 [HeLa] cells decreased, English-language ones steadily increased, from 7.7% in 2000 (1/13) to 76% (38/50) in 2017 [149]. 83.1% of HEp-2 [HeLa] literature in the MEDLINE database prior to 2018 was written by researchers in Chinese labs [149]. The Chinese Center for Type Culture Collection ("CCTCC") reported a CCCM rate of 33.6% in samples that it authenticated over the period 2010–2019 [154]. The incidence of primary CCCM dropped from 85.5% to 54.1%; the highest authenticity rate was observed in corporate scientific settings (74.7%), followed by universities (62.2%) and hospitals (51.8%) [154].

The economic stakes are underscored by comparisons to national budgets and nonprofit funding. For instance, the combined 2022 NIH budget for cancer and cancer genomics was $8.86 billion USD [155]. A conservatively estimated 22.9% rate of papers touched by CCCM-related issues means that $2.03 billion was invested in possibly-questionable research [11]. The American Cancer Society ("ACS"), the largest cancer nonprofit in the United States, distributed $145 million USD in research grants that year, making one year of potential US CCCM equivalent to nearly fourteen years of ACS awards [11,156].

Overall, these gaps in understanding and practice highlight the need for international efforts to raise awareness, inform professionals, assess adherence to guidelines, and ultimately enhance the legitimacy of global scientific research. To address this pressing issue, it is vital that the scientific community adopts more rigorous authentication methods, increases transparency, and improves the dissemination of critical information. The implementation of international standards and the strengthening of collaborative networks, such as the International Cell Line Authentication Committee ("ICLAC"), are crucial steps toward mitigating the economic and reputational consequences of CCCM.

6. Final considerations and recommendations

As we navigate the complex landscape of cellular cross-contamination and misidentification, it is clear that despite half a century of awareness, the problem persists. The experiences of pioneers like Dr. Jonas Salk and organizations such as the National Cancer Institute underscore a sobering reality: no laboratory is immune to this issue. With the rapid advancement of genetic editing technologies and the corresponding increase in engineered cell lines, the stakes are higher than ever. Authenticating these lines is critical not just for the integrity of experimental data, but also for protecting the burgeoning intellectual property they represent, which is integral to a bioeconomy projected to reach $4 trillion annually in the coming decade. The challenge of CCCM, while daunting, is not insurmountable. To make meaningful progress, a multifaceted approach is essential:

6.1. *Stricter enforcement and accountability*

- Scientific journals must enforce stringent authentication mandates. While many papers have called for this, broader adoption and stricter enforcement are imperative.
- Funding agencies should implement robust regulations to ensure that grants are allocated to projects with verified cell lines, prioritizing the reproducibility and reliability of research.

6.2. *Best practices and quality control*

- Researchers must rigorously authenticate cell lines at the start and end of each study, and regularly thereafter, ideally biannually.
- Acquisition of cell lines should be from reputable sources, with clearly documented authentication data to avoid the use of misidentified lines.
- Researchers should authenticate cell lines from collaborators upon receipt and may consult the resources presented in Table 3 while doing so.

6.3. *Community engagement and correction mechanisms*

- The community should encourage voluntary retractions and the issuance of corrigenda for published works based on misidentified cell lines, fostering a culture of accountability and transparency.
- National surveys and audits should be conducted, especially in countries with significant contributions to biological research, to assess and address the prevalence of CCCM.

6.4. *Technological innovation and data sharing*

- STR profiling remains the gold standard for cell line authentication. Creating a comprehensive database of high-loci STR profiles will enhance the effectiveness of this method, allowing for broader and more accurate screenings.
- Advanced solutions like image-based culture identification and the use of genetic Physical Unclonable Functions ("PUFs") should be explored for their potential to revolutionize cell line authentication.

The scientific community must unite to enforce global authentication standards, which will not only conserve funding but also foster the growth of the bioeconomy. By doing so, we preserve the integrity of the scientific profession and enhance the reproducibility of research, thereby reinforcing public trust in science. The concerted effort to reduce CCCM to a stringent threshold will be a testament to our commitment to scientific excellence and the unwavering pursuit of truth. As we stand on the cusp of a new era in biological research, our actions now will set a precedent for generations of scientists to come.

CRediT authorship contribution statement

Elijah Harbut: Writing – review & editing, Writing – original draft, Visualization, Investigation, Conceptualization. **Yiorgos Makris:** Writing – review & editing. **Alexander Pertsemlidis:** Writing – review & editing. **Leonidas Bleris:** Writing – review & editing, Visualization, Supervision, Funding acquisition, Conceptualization.

Table 3
CCCM resources.

CCCM Resources		
Resource	Description	Link
Organizations		
ATCC	American Type Culture Collection	https://www.atcc.org
CCTCC	China Center for Type Culture Collection	http://cctcc.whu.edu.cn
CICR	Chinese National Infrastructure of Cell Line Resource	http://www.cellresource.cn
CIMR	Coriell Institute for Medical Research	https://www.coriell.org
DSMZ	German Collection of Microorganisms and Cell Culture	https://www.dsmz.de
ECACC	European Collection of Authenticated Cell Cultures	https://www.culturecollections.org.uk
ICLAC	International Cell Line Authentication Committee	https://iclac.org
JCRB	Japanese Collection of Research Bioresources	https://cellbank.nibiohn.go.jp
NCBI	National Center for Biotechnology Information	https://www.ncbi.nlm.nih.gov
NIST	National Institute of Standards and Technology	https://www.nist.gov
NSTI	Chinese National Science and Technology Infrastructure	https://www.escience.org.cn
RBCCB	Riken Bioresource Center and Cell Bank	https://web.brc.riken.jp
WFCC	World Federation for Cell Cultures	https://wfcc.info
Databases and Tools		
ATCC STR Search	STR Profile Analysis	https://www.atcc.org/search-str-database/str-profiling-analysis
AuthentiCell	Orderable STR Kits	https://www.culturecollections.org.uk/services/authenticell
CCLA	RNA-Seq Authentication Tool	https://bioinfo.life.hust.edu.cn/web/CCLA
Cellosaurus	Cell Line Database	https://www.cellosaurus.org
CLASTR	STR Search Tool	https://www.cellosaurus.org/str-search
CLIMA	STR Search Tool	https://bioinformatics.hsanmartino.it/clima2
DSMZ	STR Search Tool	https://celldive.dsmz.de/str
ICLAC	Register of Misidentified Cell Lines	https://iclac.org/databases/cross-contaminations
NCBI Biosamples	Biological Source Material Database	https://www.ncbi.nlm.nih.gov/biosample
RRID Initiative	RRID Database	https://scicrunch.org/resources
STRBase	STR Database	https://strbase.nist.gov
Best Practices and Literature		
ANSI ASN-0002–2022	Standardized STR Protocol	https://webstore.ansi.org/standards/atcc/ansiatccasn00022022
ICLAC Guide (2014)	Guide to Human Cell Line Authentication	https://iclac.org/wp-content/uploads/Authentication-SOP_09-Jan-2014.pdf
STR Interpretation Guidelines	Guide to Cell Line Authentication by STR	https://www.ncbi.nlm.nih.gov/books/NBK144066
Bibliographic Resources	Further Reading (GLP, Cell Culturing)	[76,97,157–160][)]

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

Leonidas Bleris reports financial support was provided by The University of Texas at Dallas. Leonidas Bleris reports financial support was provided by National Science Foundation. Leonidas Bleris reports a relationship with SyntaxisBio that includes: board membership. Alex Pertsemlidis reports a relationship with SyntaxisBio that includes: board membership. Leonidas Bleris has patent Genetic Physical Unclonable Functions pending to SyntaxisBio. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We thank the Bleris lab members for support and discussions. We acknowledge funding from the US National Science Foundation (NSF) grants (2300340, 2029121, 2114192), National Institutes of Health (NIH) grant HG012884, a Cecil H. and Ida Green Endowment, and the University of Texas at Dallas.

References

[1] Culliton BJ. HeLa cells: contaminating cultures around the world. Science 1974;184(4141):1058–9 (1979).
[2] Korch CT, Capes-Davis A. The extensive and expensive impacts of HEp-2 [HeLa], intestine 407 [HeLa], and other false cell lines in journal publications. SLAS Discov 2021;12(26):1268–79.
[3] Horbach S, Halffman W. The ghosts of Hela: how cell line misidentification contaminates the scientific literature. Glanzel W, editor PLoS One 2017;10(12):e0186281.
[4] Fusenig NE, Capes-Davis A, Bianchini F, Sundell S, Lichter P. The need for a worldwide consensus for cell line authentication: experience implementing a mandatory requirement at the international journal of cancer. PLoS Biol 2017;04(15):e2001438.
[5] Dennert K, Kumar R. Traceability methods for cell line authentication and mycoplasma detection. SLAS Technol Transl Life Sci Innov 2021;26:247263032110302. 07.
[6] MacLeod RAF, Dirks WG, Matsuo Y, Kaufmann M, Milch H, Drexler HG. Widespread intraspecies cross-contamination of human tumor cell lines arising at source. Int J Cancer 1999;11(83):555–63.
[7] Nelson-Rees W, Daniels D, Flandermeyer R. Cross-contamination of cells in culture. Science 1981;04(212):446–52 (1979).
[8] Povey JF, Saintas E, Aderemi AV, Rothweiler F, Zehner R, Dirks WG, Cinatl J, Racher AJ, Wass MN. Intact-cell MALDI-ToF mass spectrometry for the authentication of drug-adapted cancer cell lines. Cells 2019;10(8):1194.
[9] Drexler HG, Dirks WG, MacLeod RAF, Uphoff CC. False and mycoplasma-contaminated leukemia-lymphoma cell lines: time for a reappraisal. Int J Cancer 2017;01(140):1209–14.
[10] Freedman LP, Gibson MC, Ethier SP, Soule HR, Neve RM, Reid YA. Reproducibility: changing the policies and culture of cell line authentication. Nat Methods 2015;12:493–7. 05.
[11] Souren NY, Fusenig NE, Heck S, Dirks WG, Capes-Davis A, Bianchini F, Plass C. Cell line authentication: a necessity for reproducible biomedical research. EMBO J 2022;06:41.
[12] Corral-Vázquez C, Aguilar-Quesada R, Catalina P, Lucena-Aguilar G, Ligero G, Miranda B, Carrillo-Ávila JA. Cell line authentication and mycoplasma detection as minimum quality control of cell lines in biobanking. Cell Tissue Bank 2017;18:271–80. 03.
[13] Fogh J, Holmgren NB, Ludovici PP. A review of cell culture contaminations. In Vitro 1971;7:26–41. 08.
[14] Kathryn Z. Points to consider in the characterization of cell lines used to produce biologicals (1993), 07. Food and Drug Administration; 1993. p. 14.
[15] Nardone RM. Curbing rampant cross-contamination and misidentification of cell lines. Biotechniques 2008;45:221–7. 09.
[16] Identity crisis. Nature 2009;457(7232):935–6.
[17] Lichter P, Allgayer H, Bartsch H, Fusenig N, Hemminki K, von Knebel Doeberitz M, Kyewski B, Miller AB. Obligation for cell line authentication: appeal for concerted action. Int J Cancer 2010;01(126):1. –1.
[18] O'Brien SJ. Cell culture forensics. Proc Natl Acad Sci 2001;07(98):7656–8.
[19] Wrigley JD, McCall EJ, Bannaghan CL, Liggins L, Kendrick C, Griffen A, Hicks R, Fröderberg-Roth L. Cell banking for pharmaceutical research. Drug Discov Today 2014;10(19):1518–29.
[20] Almeida JL, Cole KD, Plant AL. Standards for cell line authentication and beyond. PLoS Biol 2016;06(14):e1002476.
[21] Kerrigan L, Nims RW. Authentication of human cell-based products: the role of a new consensus standard. Regen Med 2011;6:255–60. 03.
[22] Masters JR, Thomson JA, Daly-Burns B, Reid YA, Dirks WG, Packer P, Toji LH, Ohno T, Tanabe H, Arlett CF, Kelland LR, Harrison M, Virmani A, Ward TH, Ayres KL, Debenham PG. Short tandem repeat profiling provides an international reference standard for human cell lines. Proc Natl Acad Sci 2001;98:8012–7. 06.
[23] Gu M, Liu J, Yang M, Zhang M, Yang J, Duan S, Ding X, Liu J, Chen C, Zeng Y, Shen C. Advantages of a 21-loci short tandem repeat method for detection of cross-contamination in human cell lines. Gene 2020;12(763):145048.
[24] Drexler HG, Dirks WG, MacLeod RAF. False human hematopoietic cell lines: cross-contaminations and misinterpretations. Leukemia 1999;10(13):1601–7.

[25] Trastulla L, Noorbakhsh J, Vazquez F, McFarland J, Iorio F. Computational estimation of quality and clinical relevance of cancer cell lines. Mol Syst Biol 2022;07:18.
[26] Pham K, Delitto D, Knowlton AE, Hartlage ER, Madhavan R, Gonzalo DH, Thomas RM, Behrns KE, George TJ, Hughes SJ, Wallet SM, Liu C, Trevino JG. Isolation of pancreatic cancer cells from a patient-derived xenograft model allows for practical expansion and preserved heterogeneity in culture. Am J Pathol 2016; 06(186):1537–46.
[27] Capes-Davis A, Theodosopoulos G, Atkin I, Drexler HG, Kohara A, MacLeod RAF, Masters JR, Nakamura Y, Reid YA, Reddel RR, Freshney RI. Check your cultures! A list of cross-contaminated or misidentified cell lines. Int J Cancer 2010;02 (127):1–8.
[28] Chang RS. HeLa marker chromosomes, chang liver cells, and liver-specific functions. Science 1978;02(199):567–8 (1979).
[29] WA NR. Response: HeLa marker chromosomes, chang liver cells, and liver-specific functions. Science 1978;02(199):567–8 (1979).
[30] O'Brien SJ, Kleiner G, Olson R, Shannon JE. Enzyme polymorphisms as genetic signatures in human cell cultures. Science 1977;195:1345–8. 03.
[31] Markovic O, Markovic N. Cell cross-contamination in cell cultures: the silent and neglected danger. In Vitro Cell Dev Biol Anim 1998;01(34):1–8.
[32] Kniss DA, Summerfield TL. Discovery of HeLa cell contamination in HES cells. Reprod Sci 2014;02(21):1015–9.
[33] Kniss DA, Xie Y, Li Y, Kumar S, Linton EA, Cohen P, Fan-Havard P, Redman CWG, Sargent IL. ED27 trophoblast-like cells isolated from first-trimester chorionic villi are genetically identical to HeLa cells yet exhibit a distinct phenotype. Placenta 2002;01(23):32–43.
[34] Boonstra JJ, van Marion R, Beer DG, Lin L, Chaves P, Ribeiro C, Pereira AD, Roque L, Darnton SJ, Altorki NK, Schrump DS, Klimstra DS, Tang LH, Eshleman JR, Alvarez H, Shimada Y, van Dekken H, Tilanus HW. Dinjens WNM. Verification and unmasking of widely used human esophageal adenocarcinoma cell lines. J Natl Cancer Inst 2010;102:271–4.
[35] Korch C, Hall EM, Dirks WG, Ewing M, Faries M, Varella-Garcia M, Robinson S, Storts D, Turner JA, Wang Y, Burnett EC, Healy L, Kniss D, Neve RM, Nims RW, Reid YA, Robinson WA. Authentication of M14 melanoma cell line proves misidentification of MDA-MB-435 breast cancer cell line. Int J Cancer 2017;10 (142):561–72.
[36] Liscovitch M, Ravid D. A case study in misidentification of cancer cell lines: MCF-7/AdrR cells (re-designated NCI/ADR-RES) are derived from OVCAR-8 human ovarian carcinoma cells. Cancer Lett 2007;01(245):350–2.
[37] Melcher R, Maisch S, Koehler S, Bauer M, Steinlein C, Schmid M, Kudlich T, Schauber J, Luehrs H, Menzel T, Scheppach W. SKY and genetic fingerprinting reveal a cross-contamination of the putative normal colon epithelial cell line NCOL-1. Cancer Genet Cytogenet 2005;158:84–7. 04.
[38] Baker M. Biotech giant publishes failures to confirm high-profile science. Nature 2016;02(530):141. –141.
[39] Begley CG, Ellis LM. Raise standards for preclinical cancer research. Nature 2012; 03(483):531–3.
[40] Baker M. 1,500 scientists lift the lid on reproducibility. Nature 2016;05(533): 452–4.
[41] Challenges in irreproducible research. Nature 2018. Springer Nature.
[42] Editors N. Reality check on reproducibility. Nature 2016;05(533):437. –437.
[43] Errington TM, Iorns E, Gunn W, Tan FE, Lomax J, Nosek BA. An open investigation of the reproducibility of cancer biology research. Elife 2014;12:3.
[44] Errington TM, Mathur M, Soderberg CK, Denis A, Perfito N, Iorns E, Nosek BA. Investigating the replicability of preclinical cancer biology. Elife 2021;12:10.
[45] Errington TM, Denis A, Perfito N, Iorns E, Nosek BA. Challenges for assessing replicability in preclinical cancer biology. Elife 2021:10.
[46] Reddin IG, Fenton TR, Wass MN, Michaelis M. Large inherent variability in data derived from highly standardised cell culture experiments. Pharmacol Res 2023; 02(188):106671.
[47] Lucey BP, WA NR, Hutchins GM. Henrietta lacks, HeLa cells, and cell culture contamination. Arch Pathol Lab Med 2009;09(133):1463–7.
[48] Gey G. Tissue culture studies of the proliferative capacity of cervical carcinoma and normal epithelium. Cancer res, 04. American Association for Cancer Research; 1952. p. 243–312.
[49] Fang Z, Lyu J, Li J, Li C, Zhang Y, Guo Y, Wang Y, Zhang Y, Chen K. Application of bioreactor technology for cell culture-based viral vaccine production: Present status and future prospects. Front Bioeng Biotechnol 2022;08:10.
[50] Cutrone R, Lednicky J, Dunn G, Rizzo P, Bocchetta M, Chumakov K, Minor P, Carbone M. Some oral poliovirus vaccines were contaminated with infectious SV40 after 1961. Cancer Res 2005;65(22):10273–9.
[51] Gold M. A conspiracy of cells: one woman's immortal legacy and the medical scandal it caused. State University Of New York Press; 1986. p. 126.
[52] Barallon R, Bauer SR, Butler J, Capes-Davis A, Dirks WG, Elmore E, Furtado M, Kline MC, Kohara A, Los GV, MacLeod RAF, Masters JRW, Nardone M, Nardone RM, Nims RW, Price PJ, Reid YA, Shewale J, Sykes G, Steuer AF, Storts DR, Thomson J, Taraporewala Z, Alston-Roberts C, Kerrigan L. Recommendation of short tandem repeat profiling for authenticating human cell lines, stem cells, and tissues. In Vitro Cell Dev Biol Anim 2010;07(46):727–32.
[53] Masters JR. HeLa cells 50 years on: the good, the bad and the ugly. Nat Rev Cancer 2002;04(2):315–9.
[54] Gartler SM. Apparent HeLa cell contamination of human heteroploid cell lines. Nature 1968;02(217):750–1.
[55] Stulberg CS, Coriell LL, Kniazeff AJ, Shannon JE. The animal cell culture collection. In Vitro 1970;12(5):1–16.
[56] Rothfels KH, Axelrad AA, Siminovitch L, McCulloch EA, Parker RC. The origin of altered cell lines from mouse, monkey and man as indicated by chromosome and transplantation studies. Can Cancer Conf 1959:189–214.
[57] Defendi V, Billingham RE, Silvers W, Moorhead P. Immunological and karyological criteria for identification of cell lines. JNCI J Natl Cancer Inst 1960; 08:25.
[58] Brand KG, Syverton JT. Results of species-specific hemagglutination tests on "transformed," nontransformed, and primary cell cultures. JNCI J Natl Cancer Inst 1962;01:28.
[59] Nelson-Rees WA. Responsibility for truth in research. Hamilton W, Weiss RA, Wain–Hobson S, editors Philos Trans R Soc Lond B Biol sci 2001;06(356):849–51.
[60] Gartler SM. Genetic markers as tracers in cell culture. Natl Cancer Inst Monogr 1967;26:167–95. 09.
[61] Nardone RM. Eradication of cross-contaminated cell lines: a call for action. Cell Biol Toxicol 2007;23(6):367–72.
[62] WA NR, Flandermeyer RR, Hawthorne PK. Banded marker chromosomes as indicators of intraspecies cellular contamination. Science 1974;06(184):1093–6 (1979).
[63] Nelson-Rees W, Flandermeyer R. Inter- and intraspecies contamination of human breast tumor cell lines HBC and BrCa5 and other cell cultures. Science 1977;03 (195):1343–4 (1979).
[64] Salk J. The spector of malignancy and criteria for cell lines as substrates for vaccines. Adv Exp Med Biol 1979;118:107–13.
[65] Lavappa KS. Survey of ATCC stocks of human cell lines for HeLa contamination. In Vitro 1978;05:469–75.
[66] Broad WJ. The case of the unmentioned malignancy. Science 1980;12(210): 1229–30 (1979).
[67] Furcinitti PS, Todd P. Gamma rays: further evidence for lack of a threshold dose for lethality to human cells. Science 1979;10(206):475–7 (1979).
[68] Harris NL, Gang DL, Quay SC, Poppema S, Zamecnik PC, Nelson-Rees WA, O'Brien SJ. Contamination of Hodgkin's disease cell cultures. Nature 1981;01 (289):228–30.
[69] Dickson D. Contaminated cell lines. Nature 1981;01(289):227. –227.
[70] Gilbert DA, Reid YA, Gail MH, Pee D, White C, Hay RJ, O'Brien SJ. Application of DNA fingerprints for cell-line individualization. Am J Hum Genet 1990;47: 499–514.
[71] Helden V, Ian ACF, Theron E, Thornley AL. Cross-contamination of human esophageal squamous carcinoma cell lines detected by DNA fingerprint analysis. Cancer Research, 10. American Association for Cancer Research; 1988. p. 5660–2.
[72] Masters JRW, Bedford P, Kearney A, Povey S, Franks LM. Bladder cancer cell line cross-contamination: identification using a locus-specific minisatellite probe. Br J Cancer 1988;57:284–6. 03.
[73] Stacey GN, Bolton BJ, Doyle A. DNA fingerprinting transforms the art of cell authentication. Nature 1992;357:261–2. 05.
[74] Ogura H, Fuji R, Hamano M, Kuzaya M, Nakajima H, Ohata R, Mori T. Detection of HeLa cell contamination-presence of human papilloma virus 18 DNA as HeLa marker in JTC-3, OG and OE cell lines. Jpn J Med Sci Biol 1997;50(4–5):161–7.
[75] Rojas A, Gonzalez I, Figueroa H. Cell line cross-contamination in biomedical research: a call to prevent unawareness. Acta Pharmacol Sin 2008;29:877–80. 07.
[76] Geraghty RJ, Capes-Davis A, Davis JM, Downward J, Freshney RI, Knezevic I, Lovell-Badge R, Masters JRW, Meredith J, Stacey GN, Thraves P, Vias M. Guidelines for the use of cell lines in biomedical research. Br J Cancer 2014;08 (111):1021–46.
[77] Hay RJ. Human cells and cell cultures: availability, authentication and future prospects. Hum Cell 1996;9:143–52. 09.
[78] Kaplan J, Hukku B. Chapter 11 cell line characterization and authentication, 01. Academic Press; 1998. p. 203–16. Mather JP, Barnes D, editors.
[79] Milanesi E, Ajmone-Marsan P, Bignotti E, Losio MN, Bernardi J, Chegdani F, Soncini M, Ferrari M. Molecular detection of cell line cross-contaminations using amplified fragment length polymorphism DNA fingerprinting technology. In Vitro Cell Dev Biol Anim 2003;39:124–30. 03.
[80] Stacey GN. Cell contamination leads to inaccurate data: we must take action now. Nature 2000;01(403):356. –356.
[81] Ruitberg CM. STRBase: a short tandem repeat DNA database for the human identity testing community. Nucleic Acids Res 2001;29:320–2. 01.
[82] Drexler HG, Dirks WG, Matsuo Y, MacLeod R a F. False leukemia-lymphoma cell lines: an update on over 500 cell lines. Leukemia 2003;02(17):416–26.
[83] Buehring GC, Eby EA, Eby MJ. Cell line cross-contamination: how aware are mammalian cell culturists of the problem and how to monitor it? In Vitro Cell Dev Biol Anim 2004;40:211.
[84] Hughes P, Marshall D, Reid Y, Parkes H, Gelber C. The costs of using unauthenticated, over-passaged cell lines: how much more data do we need? Biotechniques 2007;11(43):575–86.
[85] Dirks WG, MacLeod RAF, Nakamura Y, Kohara A, Reid Y, Milch H, Drexler HG, Mizusawa H. Cell line cross-contamination initiative: An interactive reference database of STR profiles covering common cancer cell lines. Int J Cancer 2010;01 (126):303–4.
[86] Jiang L, Zeng X, Wang Z, Chen Q. Cell line cross-contamination: KB is not an oral squamous cell carcinoma cell line. Eur J Oral Sci 2009;02(117):90–1.
[87] Capes-Davis A, Reid YA, Kline MC, Storts DR, Strauss E, Dirks WG, Drexler HG, MacLeod RAF, Sykes G, Kohara A, Nakamura Y, Elmore E, Nims RW, Alston-Roberts C, Barallon R, Los GV, Nardone RM, Price PJ, Steuer A, Thomson J, Masters JRW, Kerrigan L. Match criteria for human cell line authentication: Where do we draw the line? Int J Cancer 2012;11(132):2510–9.

[88] Vasilevsky NA, Brush MH, Paddock H, Ponting L, Tripathy SJ, LaRocca GM, Haendel MA. On the reproducibility of science: unique identification of research resources in the biomedical literature. PeerJ 2013;1:e148. 09.

[89] NOT-OD-17-068: reminder: authentication of key biological and/or chemical resources. 05 2017.

[90] Robin T, Capes-Davis A, Bairoch A. CLASTR: the cellosaurus STR similarity search tool - a precious help for cell line authentication. Int J Cancer 2019;10(146):1299–306.

[91] Bairoch A. The cellosaurus, a cell-line knowledge resource. J Biomol Tech 2018;07(29):25–38.

[92] Zhang Q, Luo M, Liu CJ, Guo AY. CCLA: an accurate method and web server for cancer cell line authentication using gene expression profiles. Brief Bioinform 2020;06:22.

[93] Almeida JL, Dakic A, Kindig K, Kone M, Letham DLD, Langdon S, Peat R, Holding-Pillai J, Hall EM, Ladd M, Shaffer MD, Berg H, Li J, Wigger G, Lund S, Steffen CR, Fransway BB, Geraghty B, Natoli M, Bauer B, Gollin SM, Lewis DW, Reid Y. Interlaboratory study to validate a STR profiling method for intraspecies identification of mouse cell lines. PLoS One 2019;06(14):e0218412.

[94] Mariyappa D., Rusch D.B., Han S., Luhur A., Overton D., Miller D.F.B., Bergman C.M., Zelhof A.C. A novel transposable element-based authentication protocol for *Drosophila* cell lines. Oliver B, editor. G3 genes|genomes|genetics. 11 2021;12.

[95] Nanda I, Schröder SK, Steinlein C, Haaf T, Buhl EM, Grimm DG, Weiskirchen R. Rat hepatic stellate cell line CFSC-2G: genetic markers and short tandem repeat profile useful for cell line authentication. Cells 2022;09(11):2900.

[96] Almeida JL, Korch CT. Authentication of human and mouse cell lines by short tandem repeat (STR) DNA genotype analysis. Eli Lilly & Company and the National Center for Advancing Translational Sciences; 2023.

[97] Nims RW, Reid Y. Best practices for authenticating cell lines. In Vitro Cell Dev Biol Anim 2017;12(53):880–7.

[98] Simpson WF, Stulberg CS, Peterson WD. Monitoring species of cells in culture by immunofluorescence. J Tissue Cult Methods 1978;4:771–4.

[99] Simpson WR, Stulberg CS. Species identification of animal cell strains by immunofluorescence. Nature 1963;08(199):616–7. Nature Portfolio.

[100] American type culture collection standards development organization workgroup ASN-0002. Cell line misidentification: the beginning of the end. Nat Rev Cancer 2010 Jun;10(6):441–8.

[101] Reid YA. Characterization and authentication of cancer cell lines: an overview. Methods Mol Biol 2011;731:35–43.

[102] Manjunath HS, James N, Mathew R, Al Hashmi M, Silcock L, Biunno I, De Blasio P, Manickam C, Tomei S. Human sample authentication in biomedical research: comparison of two platforms. Sci Rep 2021;11:13982. 07.

[103] Lung O, Candlish R, Nebroski M, Kruckiewicz P, Buchanan C, Moniwa M. High-throughput sequencing for species authentication and contamination detection of 63 cell lines. Sci Rep 2021;11:11.

[104] Huang Y, Liu Y, Zheng C, Shen C. Investigation of cross-contamination and misidentification of 278 widely used tumor cell lines. Talamas-Rohana P, editor PLoS One 2017;01(12):e0170384.

[105] Hebert PDN, Cywinska A, Ball SL, deWaard JR. Biological identifications through DNA barcodes. Proc Biol Sci 2003;270:313–21.

[106] Didion JP, Buus RJ, Naghashfar Z, Threadgill DW, Morse HC, de Villena FPM. SNP array profiling of mouse cell lines identifies their strains of origin and reveals cross-contamination and widespread aneuploidy. BMC Genomics 2014;15:847.

[107] Cooper JK, Sykes G, King S, Cottrill K, Ivanova NV, Hanner R, Ikonomi P. Species identification in cell culture: a two-pronged molecular approach. In Vitro Cell Dev Biol Anim 2007;10(43):344–51.

[108] Castro F, Dirks WG, Fähnrich S, Hotz-Wagenblatt A, Pawlita M, Schmitt M. High-throughput SNP-based authentication of human cell lines. Int J Cancer 2012;06(132):308–14.

[109] Dittmar KEJ, Simann M, Zghoul N, Schön O, Meyring W, Hannig H, Macke L, Dirks WG, Miller K, Garritsen HSP, Lindenmaier W. Quality of cell products: authenticity, identity, genomic stability and status of differentiation. Transfus Med Hemother 2010;37:2. –2.

[110] Yan R, Ottenbreit MJ, Hukku B, Mally M, Chou S, Kaplan J. DNA fingerprinting of human cell lines using PCR amplification of fragment length polymorphisms. In Vitro Cell Dev Biol Anim 1996;11(32):656–62. Springer Science+Business Media.

[111] Reid Y, Storts D, Riss T, Minor L. Authentication of human cell lines by STR DNA profiling analysis, 05. Eli Lilly & Company and the National Center for Advancing Translational Sciences; 2013.

[112] Oei RW, Hou G, Liu F, Zhong J, Zhang J, An Z, Xu L, Yang Y. Convolutional neural network for cell classification using microscope images of intracellular actin networks. Horvath D, editor PLoS One 2019;14:e0213626. 03.

[113] Mzurikwao D, Khan MU, Samuel OW, Cinatl J, Wass M, Michaelis M, Marcelli G, Ang CS. Towards image-based cancer cell lines authentication using deep neural networks. Sci Rep 2020;11(10):19857.

[114] Liu C, Wang Z, Jia J, Liu Q, Su X. High-content video flow cytometry with digital cell filtering for label-free cell classification by machine learning. Cytometry A 2023 Apr;103(4):325–34.

[115] Gu Y, Zhang AC, Han Y, Li J, Chen C, Lo YH. Machine learning based real-time image-guided cell sorting and classification. Cytometry A 2019;95(5):499–509. May.

[116] Xu J, Zhou D, Deng D, Li J, Chen C, Liao X, Chen G, Heng PA. Deep learning in cell image analysis. Intell Comput 2022. 2022.

[117] Durkee MS, Abraham R, Clark MR, Giger ML. Artificial intelligence and cellular segmentation in tissue microscopy images. Am J Pathol 2021;191(10):1693–701.

[118] Chen J, Xu L, Li X, Park S. Deep learning models for cancer stem cell detection: a brief review. Front Immunol 2023;14:1214425.

[119] Yang X, Chen D, Sun Q, Wang Y, Xia Y, Yang J, Lin C, Dang X, Cen Z, Liang D, Wei R, Xu Z, Xi G, Xue G, Ye C, Wang LP, Zou P, Wang SQ, Rivera-Fuentes P, Püntener S, Chen Z, Liu Y, Zhang J, Zhao Y. A live-cell image-based machine learning strategy for reducing variability in PSC differentiation systems. Cell Discov 2023;9(1):53. Jun 6.

[120] Cheuque C, Querales M, León R, Salas R, Torres R. An efficient multi-level convolutional neural network approach for white blood cells classification. Diagnostics 2022;12(2) (Basel)Jan 20.

[121] Nissim N, Dudaie M, Barnea I, Shaked NT. Real-time stain-free classification of cancer cells and blood cells using interferometric phase microscopy and machine learning. Cytometry A 2021;99(5):511–23. May.

[122] Li Y, Zaheri S, Nguyen K, Liu L, Hassanipour F, Bleris L. Machine learning-based approaches for identifying human blood cells harboring CRISPR-mediated fetal chromatin domain ablations. Sci Rep 2022;12(1):1481. Jan 27.

[123] Li Y, Nowak CM, Pham U, Nguyen K, Bleris L. Cell morphology-based machine learning models for human cell state classification. NPJ Syst Biol Appl 2021;7(1):23.

[124] Bian X, Yang Z, Feng H, Sun H, Liu Y. A Combination of species identification and STR profiling identifies cross-contaminated cells from 482 human tumor cell lines. Sci Rep 2017;08:7.

[125] Msalbi D, Elloumi-Mseddi J, Hakim B, Sahli E, Aifa S. Chemical alternative for cell identification and cross-contamination detection. 3 Biotech 2022;12. 02.

[126] Bai H, Liu Z, Zhang TF, Du J, Zhou C, He W, Joe R, Jacky TBZ. Multifunctional supramolecular assemblies with aggregation-induced emission (AIE) for cell line identification, cell contamination evaluation, and cancer cell discrimination. ACS Nano, 06. American Chemical Society; 2020. p. 7552–63.

[127] Fasterius E, Raso C, Kennedy S, Rauch N, Lundin P, Kolch W, Uhlén M, Al-Khalili Szigyarto C. A novel RNA sequencing data analysis method for cell line authentication. Costa M, editor PLoS One 2017;02(12):e0171435.

[128] Mohammad TA, Tsai YS, Ameer S, Chen HIH, Chiu YC, Chen Y. CeL-ID: cell line identification using RNA-seq data. BMC Genomics 2019;02:20.

[129] Bandrowski A, Brush M, Grethe JS, Haendel MA, Kennedy DN, Hill S, Hof PR, Martone ME, Pols M, Tan S, Washington N, Zudilova-Seinstra E, Vasilevsky N. The resource identification initiative: a cultural shift in publishing. F1000Res 2015;11(4):134.

[130] Babic Z, Capes-Davis A, Martone ME, Bairoch A, Ozyurt IB, Gillespie TH, Bandrowski AE. Incidences of problematic cell lines are lower in papers that use RRIDs to identify cell lines. Elife 2019;01:8.

[131] Plant AL, Halter MW, Stinson JR, Greene GR. Contemporaneous sample data tracking for the generation of genome edited cell lines. Sci Rep 2022;12(12).

[132] Tellechea-Luzardo J, Hobbs L, Velázquez E, Pelechova L, Woods S, de Lorenzo V, Krasnogor N. Versioning biological cells for trustworthy cell engineering. Nat Commun 2022;02(13):765.

[133] Tellechea-Luzardo J, Winterhalter C, Widera P, Kozyra J, de Lorenzo V, Krasnogor N. Linking engineered cells to their digital twins: a version control system for strain engineering. ACS Synth Biol 2020;02(9):536–45.

[134] Chen X, Qian W, Song Z, Li QX, Guo S. Authentication, characterization and contamination detection of cell lines, xenografts and organoids by barcode deep NGS sequencing. NAR Genom Bioinform 2020;08:2.

[135] Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. Science 2012;337(6096):816–21 (1979).

[136] Cong L, Ran FA, Cox D, Lin S, Barretto R, Habib N, Hsu PD, Wu X, Jiang W, Marraffini LA, Zhang F. Multiplex genome engineering using CRISPR/Cas systems. Science 2013;339(6121):819–23 (1979).

[137] Mali P, Yang L, Esvelt KM, Aach J, Guell M, DiCarlo JE, Norville JE, Church GM. RNA-guided human genome engineering via Cas9. Science 2013;339(6121):823–6 (1979).

[138] Nowak CM, Lawson S, Zerez M, Bleris L. Guide RNA engineering for versatile Cas9 functionality. Nucleic Acids Res 2016;44(20):9555–64.

[139] Lawson S, Bleris L. Techniques and strategies employing engineered transcription factors. Curr Opin Biomed Eng 2017;4:152–62.

[140] Li Y, Nowak CM, Withers D, Pertsemlidis A, Bleris L. CRISPR-based editing reveals edge-specific effects in biological networks. CRISPR J 2018;1:286–93.

[141] Martin DK, Vicente O, Beccari T, Kellermayer M, Koller M, Lal R, Marks RS, Marova I, Mechler A, Tapaloaga D, Žnidaršič-Plazl P, Dundar M. A brief overview of global biotechnology. Biotechnol Biotechnol Equip 2021;35:S5–14. 02.

[142] National Academies of Sciences. Engineering, and medicine; division on engineering and physical sciences; health and medicine division; policy and global affairs; division on earth and life studies; forum on cyber resilience; board on health sciences policy; board on science, technology, and economic policy; board on agriculture and natural resources; board on life sciences; committee on safeguarding the bioeconomy: finding strategies for understanding, evaluating, and protecting the bioeconomy while sustaining innovation and growth. Safeguarding the bioeconomy, 03. WashingtonDC: National Academies Press (US); 2020.

[143] Mao N, Aggarwal N, Poh CL, Cho BK, Kondo A, Liu C, Yew WS, Chang MW. Future trends in synthetic biology in Asia. Adv Genet 2021 Mar;2(1):e10038.

[144] Gallo M. The bioeconomy: a primer. Congr Res Serv 2022;09.

[145] Huang K, Carulli JM, Makris Y. Counterfeit electronics: a rising threat in the semiconductor manufacturing industry. In: Proceedings of the 2013 IEEE international test conference (ITC); 2013. p. 1–4.

[146] Lofstrom K, Daasch WR, Taylor D. IC identification circuit using device mismatch. In: Proceedings of the 2000 IEEE international solid-state circuits conference digest of technical papers (Cat No00CH37056); 2000. p. 372–3.

[147] Li Y, Bidmeshki MM, Kang T, Nowak CM, Makris Y, Bleris L. Genetic physical unclonable functions in human cells. Sci Adv 2022;05:8.
[148] Freedman LP, Cockburn IM, Simcoe TS. The economics of reproducibility in preclinical research. PLoS Biol 2015;06(13):e1002165.
[149] Gorphe P. A comprehensive review of Hep-2 cell line in translational research for laryngeal cancer. Am J Cancer Res 2019;9:644–9. e-Century Publishing Corporation;.
[150] Tollefson J. China declared world's largest producer of scientific articles. Nature 2018;01(553):390. –390.
[151] China: Agriculture key component of bio-economy Five-Year Plan. USDA foreign agricultural service. 05 2022.
[152] China's five-year bioeconomy plan to focus on low-carbon growth, epidemic prevention. 05 2022.
[153] Ye F, Chen C, Qin J, Liu J, Zheng C. Genetic profiling reveals an alarming rate of cross-contamination among human cell lines used in China. FASEB J 2015;10(29):4268–72.
[154] Gu M, Yang M, He J, Xia S, Zhang Z, Wang Y, Zheng C, Shen C. A silver lining in cell line authentication: Short tandem repeat analysis of 1373 cases in China from 2010 to 2019. Int J Cancer 2021;150:502–8. 09.
[155] National Institutes of Health. Funding reports. report.nih.gov; 2020.
[156] How American cancer society research funding works. www.cancer.org.
[157] Carrillo-Ávila JA, Catalina P, Aguilar-Quesada R. Quality control of cell lines using DNA as target. DNA 2022;02(2):44–55 (Basel).
[158] Roth JS, Lee TD, Cheff DM, Gosztyla ML, Asawa RR, Danchik C, Michael S, Simeonov A, Klumpp-Thomas C, Wilson KM, Hall MD. Keeping it clean: the cell culture quality control experience at the national center for advancing translational sciences. SLAS DISCOV Adv Sci Drug Discov 2020;25:491–7. 04.
[159] Vicente E, Lesniewski M, Newman D, Vujaskovic Z, Jackson IL. Best practices for authentication of cell lines to ensure data reproducibility and integrity. Radiat Res 2021;12.
[160] Weiskirchen S, Schröder SK, Buhl EM, Weiskirchen R. A beginner's guide to cell culture: practical advice for preventing needless problems. Cells 2023;12:682. 01.