

40nm Embedded Self-Aligned Split-Gate Flash Technology for High-Density Automotive Microcontrollers

L.Q. Luo, Y.J. Kong, F.X. Deng, X. Qu, Z.Q. Teo, J.Q. Liu, F. Zhang, X.S. Cai, K.M. Tan, K.Y. Lim, P. Khoo, P.Y. Yeo, B.Y. Nguyen, S.M. Jung, S.Y. Siah, D. Shum*, Sr. Member, IEEE

GLOBALFOUNDRIES Singapore Pte, Ltd., Singapore, 738406.

K.L. Pey, K. Shubhakar, L.Q. Luo

Singapore University of Technology and Design, 8 Somapah Road, Singapore, 487372.

C.M. Wang, J.C. Xing, G.Y. Liu, Y. Diao, G.M. Lin, F. Luo, L. Tee, V. Markov, S.M. Lemke, P. Ghazavi, N. Do, V. Tiwari, X. Liu, Sr. Member, IEEE

Silicon Storage Technology, Inc., A Subsidiary of Microchip Technology Inc., San Jose, CA 95134, U.S.A.

8th International Memory Workshop
Monterey, CA, USA
May 2017



40nm Embedded Self-Aligned Split-Gate Flash Technology for High-Density Automotive Microcontrollers

L.Q. Luo, Y.J. Kong, F.X. Deng, X. Qu, Z.Q. Teo, J.Q. Liu, F. Zhang, X.S. Cai, K.M. Tan, K.Y. Lim, P. Khoo, P.Y. Yeo, B.Y. Nguyen, S.M. Jung, S.Y. Siah, D. Shum, Sr. Member, IEEE

GLOBALFOUNDRIES Singapore Pte, Ltd., Singapore, 738406.

K.L. Pey, K. Shubhakar, L.Q. Luo

Singapore University of Technology and Design, 8 Somapah Road, Singapore, 487372.

C.M. Wang, J.C. Xing, G.Y. Liu, Y. Diao, G.M. Lin, F. Luo, L. Tee, V. Markov, S.M. Lemke, P. Ghazavi, N. Do, V. Tiwari, X. Liu, Sr. Member, IEEE

Silicon Storage Technology, Inc., A Subsidiary of Microchip Technology Inc., San Jose, CA 95134, U.S.A.

Abstract—This paper successfully demonstrates a logic-compatible, high performance and high reliability, automotive-grade 2.5V embedded NVM process extending over several generations. A high-density flash macro is used to debug process complexities which arise from the add-on modules. The modular approach is adopted for integrating self-aligned, floating-gate-based split-gate SuperFlash[®] ESF3 cell into 40nm CMOS logic process. Key features of the product-like Macro are dual power supply with input voltage fluctuations, wide operating temperature range from -40°C to 150°C, fast byte/word program under 10 μ s and sector/chip erase under 10ms. The macro random read access time is only 8ns under worst case conditions. Key process monitors are characterization and yield of the Macro. Endurance was extended to 200k cycles and satisfy automotive grade requirement with wide read margin. Post-cycling data retention performs very well up to 150°C. Wafer sort yield is in high double digits, with consistent wafer-to-wafer and within-wafer uniformity, showing good process control. The technology is suitable for high-speed automotive MCU, as well as IoT, smart card, and industrial MCU applications.

Introduction

With the growth of solid-state non-volatile memories (NVM) in storing configuration settings, program code, application parameters and data in consumer, industrial and automotive electronics, demand for embedded NVM is rapidly increasing. The competition in compact product design is driving the embedded Flash memory technology to its scaling limit [1-2]. For automotive applications, robust quality and reliability are required, with zero failure rates at harsh temperature conditions. Code storage requires 10 ns fast random access, 1K endurance, >10

years of retention, while data storage requires 200K endurance, 10 years retention but slower access. The ESF3 cell uses highly efficient Source-Side CHE program, poly-to-poly FN tunneling erase, and low-voltage read, with unique advantages in providing over-erase immunity [3]. It offers simple design architecture, high speed and low power operation, good manufacturability and high reliability [4]. It ideally serves the automotive as well as general MCU and smart card markets. It is crucial to follow the promising scaling path from previous ESF3 generations [5]. The 40 nm cell size was reported previously, characterized using 16Mb Design test Chip (DTC) array [6]. In this paper we successfully demonstrate a functionally competitive, logic-compatible, automotive-grade embedded process with a high degree of modularity, using 20Mb product-like Flash Macro, together with a 32Mb SRAM macro fabricated on the same silicon, to debug any process complexities that arise from the add-on modules. We present read shmoos, power consumption, array characterization, and endurance data at 25°C and 150°C.

Process Integration

The ESF3 memory cell is erased via poly-to-poly FN tunneling across a field-enhanced asymmetric tunneling barrier formed between the floating gate (FG) and erase gate (EG) [7] and programmed via Source-side CHE Injection through the FG oxide. Compared to other split-gate (SG) FG cell concepts which program and erase through the same tunnel dielectric, and charge-trapping cells which rely on hot carrier injection with non-overlapping injection

points, the ESF3 cell is naturally less susceptible to P/E cycling degradations. Read voltage is limited to dual power supply ranges, without the need for voltage pumping. This unique combination of operating conditions enables high speed and low power operation required in automotive applications.

Fig. 1 shows how the embedded NVM (Flash memory, High Voltage (HV) peripheral) modules are integrated into the 40nm logic baseline process flow. Fig. 2 shows the Flash cell structures with a TEM cross-section along the cell channel. A successful embedded NVM process must carefully control the impact of additional process steps (thermal and wet) on baseline logic devices. We choose to integrate NVM process steps before the logic module. First module after wells formation is the FG oxide through which hot electrons are injected during programming. Next is the FG polysilicon formation, followed by ONO, Control Gate (CG) polysilicon deposition, and CG/ONO/FG stack patterning. Several dielectric spacers are formed during this process to provide robust isolation between adjacent nodes and to create optimized geometries for enhancing electric fields during erase, program and read. The HV oxide, tunnel oxide, IO gate oxide and core logic gate oxide are formed before deposition of logic polysilicon which is shared by logic, IO, cell WL, cell EG, and HV gates. HV LDD is implanted after the gate is patterned. After this step, the whole NVM and HV module are integrated into the logic. The rest of the steps follow a logic baseline process.

Results and Discussion

Flash Array Operating Condition summarized in Table 1.

Table 1 <i>ESF3-55/40</i>	<i>Erase</i>	<i>Program</i>	<i>Read</i>
Word-Line (WL)	0V	~ 1V	Vdd
Bit-Line (BL)	0V	1 μ A	0.6-1.1V
Coupling Gate (CG)	0V	10.5V	Vdd
Erase Gate (EG)	12V	4.5V	0V
Source-Line (SL)	0V	4.5V	0V

Key automotive parameters of the 20Mb Flash Array data match well with previously reported chip simulation [6], and have shown high yield consistently in high double digits, with good wafer-to-wafer and within-wafer uniformity, which indicates good process control. Key design specs are summarized in Table 2, with Die photo in Fig. 3.

Table 2 <i>ATV Macro</i>	<i>Flash Key Specs</i>	<i>Silicon data</i> [#]
Memory density	20Mb	
Sector Size	256 x 38 bits	
Dual Power Supply	1.1V \pm 10%; 2.5V \pm 10%	
Operation Temperature Tj	-40°C to 150°C	

Word Program Time	10 μ s	< 10 μ s
Read access Time	random 10 ns	~ 8 ns
Sector Erase Time	10ms	< 10ms
Chip Erase Time	20ms	< 20ms
Wake-up time	5 μ s from deep power down	pass
P/E Endurance	\geq 100K w/o ECC \geq 200K w/ ECC	Pass
Data Retention	> 10 yrs @125°C	Pass

[#]typical condition: Vdd=2.25V~2.75V; Vdd11=0.99V~1.21V; temperature: -40°C ~ 150°C.

Fig. 4 shows random read access Shmoo plot, demonstrating one of the fastest 40 nm e-NVM chips in the world, with random access time under 8 ns at 25°C and 8.1 ns at 150°C. Fig. 4 shows a wide passing window under both Vdd \pm 10% and Vdd11 \pm 10%.

Active/Standby Power Consumptions for Automotive 20Mb Flash Macro in Typical Condition (TT at 25°C, Vdd11, Vdd 2.5V) are summarized in Table 3, showing outperforming performance.

Table 3 Category	Silicon Data	Unit
Read	53	μ A/MHz
Program	1.8	mA
Sector Erase	0.6	mA
Standby	285	μ A
Deep Power Down	44	μ A

Array Characterization Data for 20Mb Automotive Macro at 25°C and 150°C are shown below. Read/Program/ Erase characteristics are shown in Fig. 5: Cell read current has tight distribution in both on- and off-states, with good margin to read reference level (Fig. 5a); Erase Voltage has tight distribution with good margin to standard user erase condition (Fig. 5b); and programmed state CGVT has tight distribution with good margin to standard user read condition (Fig. 5c). Program Disturb window is sufficient. No read failures were observed in the 20Mb array after 20 times program FF stress cycles at temperatures up to 150°C.

Fig. 6 shows endurance capability under continuous P/E cycling at 25°C and 150°C. Endurance Life is capable of \geq 100K without using error correction, and \geq 200K using Single Error Correction. Data Retention lifetime is > 10 years @ 125°C, as previously reported [6]. Overall eFlash

performance and reliability are sufficient for meeting the market demands for next-generation automotive MCUs.

For any embedded Flash integration into a logic process, logic device performance must remain unchanged. It is vital to check logic and SRAM devices' performance to identify any potential transistor degradation. The final in-line parameters show that, with additional eFlash process, the logic transistor performance and centering are closely matched to baseline. 32Mb high-density POR SRAM macros have been included in the development. Fig. 7 shows that the SRAM chip is fully functional at both V_{max} and V_{min} with comparable macro yield. The majority failures are SB, indicating that the process has no intrinsic issues and shows no major roadblocks to the overall embedded yield. The concept of integrating the Flash modules prior to the logic well has been verified.

Conclusion

In this work we have successfully integrated a functional 2.5V ESF3 Array into a 40nm CMOS logic process with copper and low-K interconnects. This highly modular integration has been verified and match SRAM performance, which makes it ideal for potential combinations with mmWave or Analog/power for single chip solutions. High density up to 20Mb product-like automotive array functionality and tight V_t /current distributions are demonstrated. SSI programming, poly-to-poly FN erase, 2.5V read, and under 10ns random read access allow for low-power, high-speed operation, suitable for automotive, general purpose MCU, and smart card markets.

References

- [1] Y.S. Shin et al., VLSI Symp. Tech. Dig., 2005, pp. 156-159J.
- [2] D. Shum et al., IMW, pp.139-142, 2012; R. Strenz, IEDM 2011, pp. 211-214.
- [3] B.J.Ahn et al., IEEE EDL, vol. 19, pp. 438-440, Nov. 1998; S. Kianian et al., Symp. VLSI Tech. Dig., pp.71-72 (1994); R. Mih et al, VLSI Symp. Tech. Dig., 2000, pp. 120-121.
- [4] Y.H. Wang et al., IEEE TED, vol. 52, pp. 385-391, Mar 2005; W.T. Chu et a., IEEE TED, vol 51, pp. 1667-1671, Oct 2004; Y. Taito et al., ISSCC 2015, pp. 132-133.
- [5] N. Akil et al., IEEE TED, vol52, pp. 492-499, Apr il 2005.
- [6] L.Q. Luo et al., IMW, pp.165-168, 2016.
- [7] Tkachev, et al., IEEE TED,, vol 59, issue 1, pp.5, 2012.

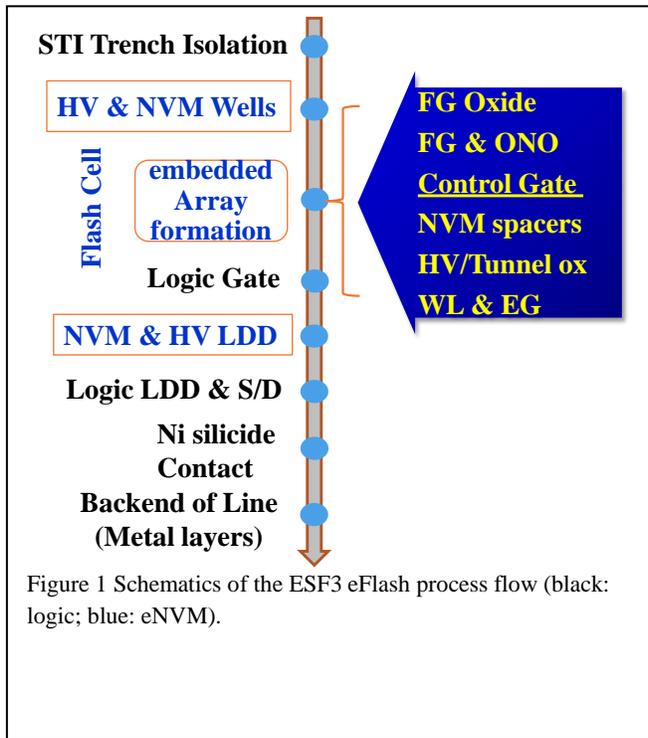


Figure 1 Schematics of the ESF3 eFlash process flow (black: logic; blue: eNVM).

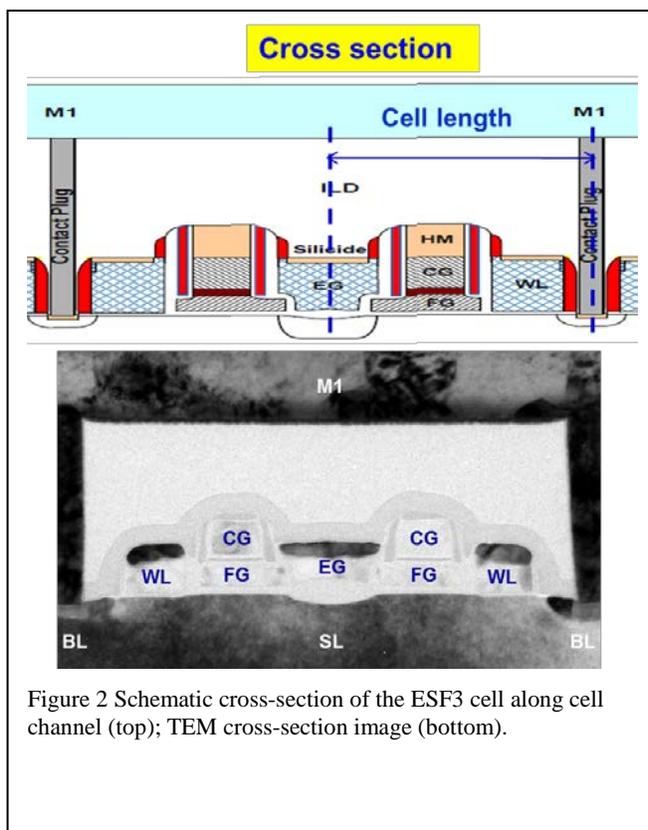


Figure 2 Schematic cross-section of the ESF3 cell along cell channel (top); TEM cross-section image (bottom).

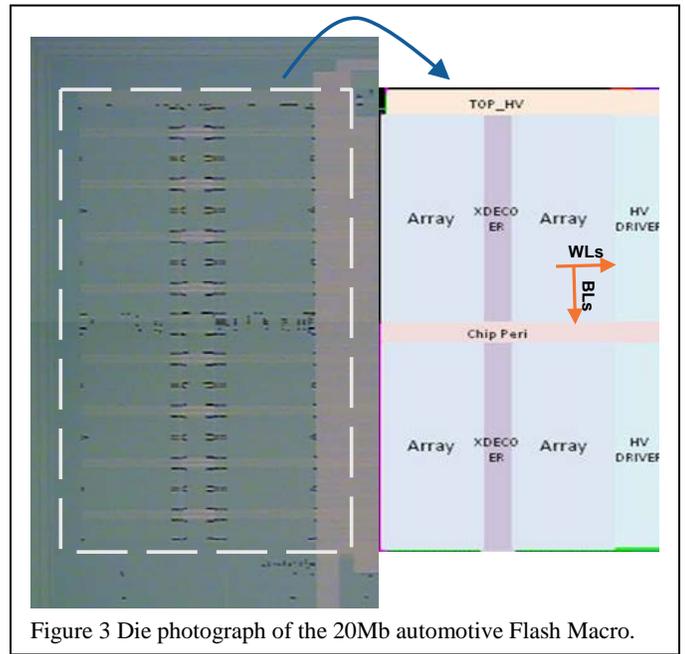


Figure 3 Die photograph of the 20Mb automotive Flash Macro.

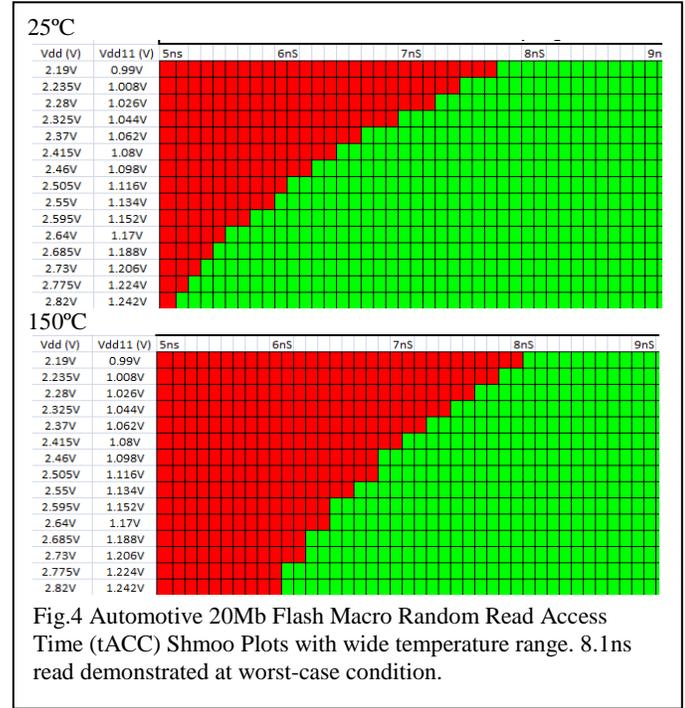


Fig.4 Automotive 20Mb Flash Macro Random Read Access Time (tACC) Shmoo Plots with wide temperature range. 8.1ns read demonstrated at worst-case condition.

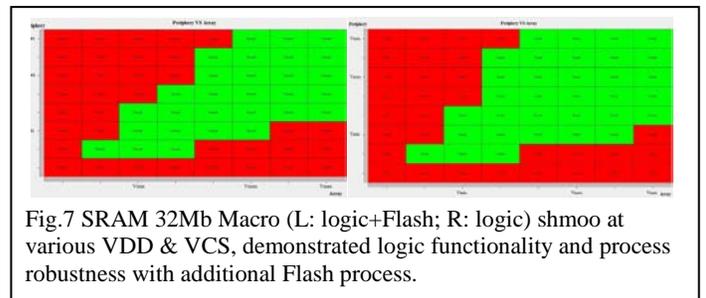


Fig.7 SRAM 32Mb Macro (L: logic+Flash; R: logic) shmoo at various VDD & VCS, demonstrated logic functionality and process robustness with additional Flash process.

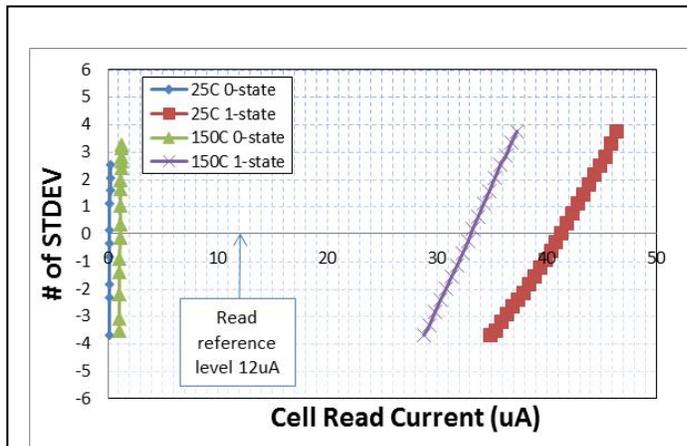


Fig. 5a) Tight distribution in both on- and off-state Cell read current I_{rl} , with good margin to read reference level.

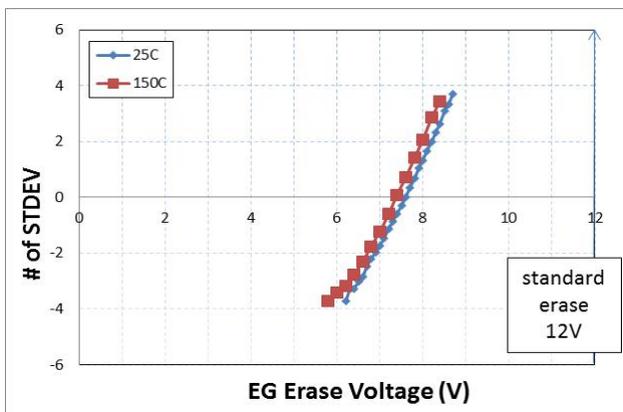


Fig. 5b) Tight erase distribution with good margin to standard user erase condition.

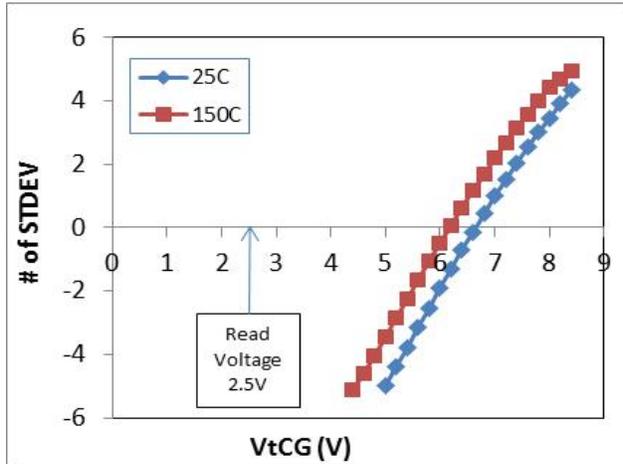


Fig. 5c) Tight program distribution with good margin to standard user read condition.

SuperFlash® is a registered trademark of Microchip Technology Incorporated in the U.S.A. and other countries. All other trademarks mentioned herein are the property of their respective companies.

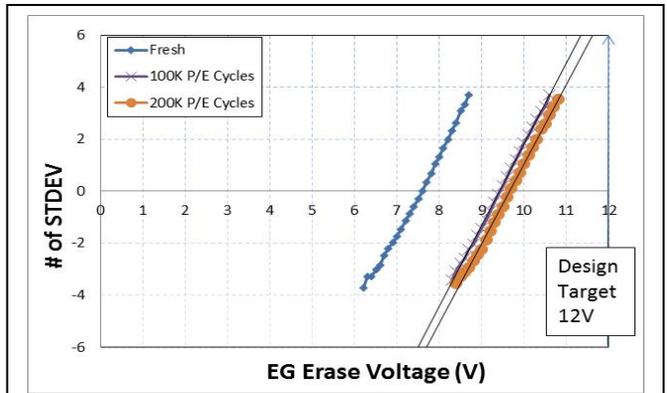


Fig. 6a 25C Erase V distribution with good margin after cycling.

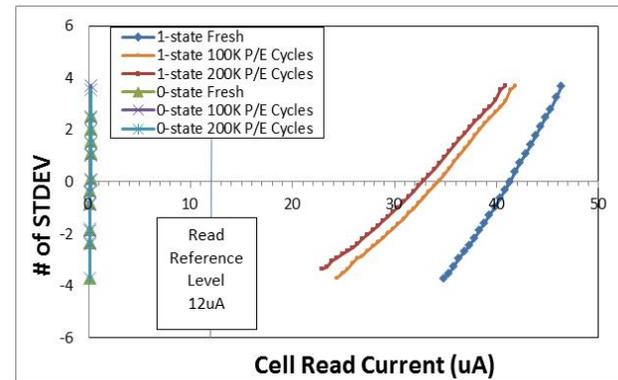


Fig. 6b 25C cell read current distribution, with good margin to read reference level.

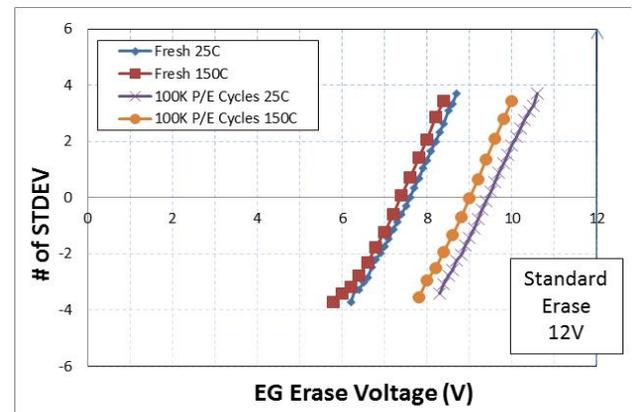


Fig. 6c Erase voltage distribution for 25C cycling vs. 150C cycling.

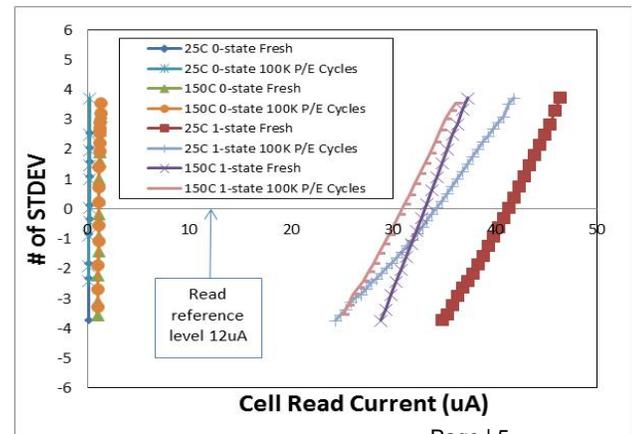


Fig. 6d Read current distribution for 25C cycling vs. 150C cycling.